

**ANALYSIS OF THE PSYCHOMETRIC PROPERTIES OF THE ZAMBIA
ACHIEVEMENT TEST: READING RECOGNITION**

Thesis
M.ED.
(PSY)
Cha
2006

By

Florence Chamvu

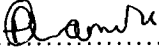
A dissertation submitted in partial fulfilment of the requirements of the degree of
Master of Education in Educational Psychology.

The University of Zambia
2006



AUTHOR'S DECLARATION

I, Florence Chamvu do solemnly declare that this dissertation is my own work and that it has not previously been submitted for a degree at this or any other university.

Signature: 

Date: 07/06/06

APPROVAL

This dissertation of Florence Chamvu is approved as fulfilling part of the requirements for the award of the degree of Master of Education in Educational Psychology by the University of Zambia.

Signed:.....


Date: 07/08/06.....

Signed:.....


Date: 07/06/06.....

Signed:.....

Date:.....

Signed:.....

Date:.....

NOTICE OF COPYRIGHT

All rights reserved. No part of this dissertation may be reproduced, stored in any retrieval form or transmitted in any form or by any means such as photocopying without the written permission of the author.

ABSTRACT

This study set out to analyze the psychometric properties of the Reading Recognition part of the Zambia Achievement Test. This test was developed by a team of researchers at the Yale University and was adapted to the Zambian situation by the University of Zambia Researchers. The test was then translated into Chinyanja. The main purposes of this study were: (1) to analyse properties of the two versions of the test and: (2) to find out if the two tests were comparable in terms of their ability to discriminate between good and bad readers.

A sample of 240 children from grades 4 to 6 were assessed, 121 were assessed in English and 119 were assessed in Chinyanja. The two groups of children were drawn from two schools in Lusaka, Ng'ombe Basic School and Olympia Basic School. The pupils were assigned to either the Chinyanja condition or the English condition depending on which language they spoke at home and were most comfortable with. The pupils were taken through training instructions to make sure that they understood the task. The test, which consisted of letter-matching, sound-picture matching letter-sound matching and word identification, was given to the pupils individually. There were no stoppage rules and all the pupils had to try all the items in the test. The responses were recorded on a bubble sheet.

The results from the two groups of children were subjected to several levels of analysis. Item difficulty and discrimination, reliability and the means and standard deviations were all computed. T – tests were applied at each grade level to compare the reading achievement in the two groups of children. Analysis of variance was also applied. Results were as follows:

- The Chinyanja test had poor discrimination ability at among good readers.
- The reliability of both versions of the test was very high (alpha = 0.98 for Chinyanja and alpha = 0.97 for English).
- The Chinyanja group performed significantly better than English group on the test as a whole.
- The English group was significantly superior to the Chinyanja group on the phoneme identification tasks (sound-picture matching and letter sound matching and sound-word matching).
- A comparison of the means at grade level revealed that in grade 5, there was no difference at all in the mean scores.
- Grade 6 results showed that although the Chinyanja group was ahead of the English group but the difference was insignificant. At grade 4 level the Chinyanja group of pupils scored significantly higher than the English group.

The difference in the performance of the two groups was interpreted as being due to the differences in the writing systems of the two languages. Pupils learn to read faster in Chinyanja than in English because the Chinyanja writing system is transparent or regular. Children in the Chinyanja group performed poorly on the phoneme identification tasks because the instructions were too complicated for them. This study has illustrated some of the problems that researchers face when using translating translated tests. As both version of the test are very reliable and the validity is good they can therefore be useful in assessing children in primary schools when the few recommended changes have been made.

ACKNOWLEDGEMENT

This dissertation would not have been possible without the help of many people. Many thanks to my supervisors Dr D. M. Kalabula and Mr Aidan Mambwe for their unrelenting patience in correcting me and reading my several revisions. I would like to thank Professor Robert Serpell for his very valuable guidance and support.. I would also like to thank Professor Lyytinen of the University of Jyvaskyla and Mikko Aro of Niilo Maki Institute for their inspiration and for providing me with the necessary literature for my research.

TABLE OF CONTENTS

Title page.....	i
Declaration	ii
Notice of copyright.....	iii
Approval.....	iv
Abstract	v
Acknowledgement.....	vii
Table of contents.....	viii
List of figures.....	xi
List of tables.....	xii
Chapter One: Introduction.....	1
1.1 Background	1
1.1.1 Knowledge relevant for reading.....	3
1.1.2 The English writing system.....	5
1.1.3 The Chinyanja orthography.....	6
1.2 Statement of the problem.....	8
1.3 Purpose of the study.....	9
1.4 Specific objectives.....	9
1.5 Research questions.....	9
1.6 Significance of the study.....	10
1.7 Limitations	11
1.8 Operational definitions.....	11
1.9 Theoretical framework.....	14
Chapter Two: Literature Review.....	17
2.1 Letter knowledge and reading.....	18
2.2 Phonological awareness and reading.....	19

2.3 Analysis of reading tests.....	21
2.4 Problems of translation.....	22
2.5 Effects of orthographies on reading.....	24
2.6 Reading studies carried out in Zambia.....	27
2.7 Zambian languages in the classroom.....	33
2.8 Reliability.....	36
2.9 Validity.....	37
2.10 Summary of reviewed literature.....	39
Chapter Three: Methodology.....	41
3.1 The sample.....	41
3.2 Instruments.....	42
3.3 Procedure	42
3.4 Data analysis.....	43
Chapter Four: Results.....	45
4.1 Item difficulty and discrimination indexes.....	45
4.2 Reliability results	47
4.3 Factor analysis results.....	49
4.4 Distracter analysis results.....	49
4.5 Means, standard deviations and t-values.....	50
Chapter Five: Discussion.....	64
5.1 Item difficulty and item discrimination.....	64
5.2 Reliability.....	66
5.3 Quality of distracters.....	70
5.4 Construct validity.....	70
5.5 Means.....	71

5.6 Overall performance.....	78
Chapter Six: Conclusion.....	83
6.1 Summary of results.....	83
6.2 Conclusion.....	85
6.2 Recommendations.....	85
References.....	86
Appendices	
Appendix A: Demographic data sheet.....	92
Appendix B: Sample of items in the ZAT-RR.....	93
Appendix C: Item difficulty and discrimination indexes.....	94

LIST OF FIGURES

Figure I Grade 4 English histogram.....51

Figure II Grade 4 Nyanja histogram.....52

Figure III Grade 5 Nyanja histogram.....54

Figure IV Grade 5 English histogram.....55

Figure V Grade 6 Nyanja histogram.....57

Figure VI Grade 6 English histogram.....58

Figure VII Distribution of scores for the English group.....60

Figure VIII Distribution of scores for the Nyanja group.....61

LIST OF TABLES

Table 1 Grade 4 means standard deviations and t-values.....53

Table 2 Grade 5 means standard deviations and t-values.....56

Table 3 Grade 6 means standard deviations and t-values.....59

Table 4 ANOVA results.....62

CHAPTER ONE: INTRODUCTION

A number of psychological tests are used in Zambia for educational purposes. The results of these tests often influence decisions in educational settings. A considerable number of these tests were imported from abroad. Unfortunately, the psychometric characteristics of many of these tests, including some of the locally constructed ones, are not reported. A survey done by Machungwa and Ali (1984) showed that there were fourteen locally developed and imported educational tests being used and most of these tests either did not have statistical information or the psychometric properties were reported to be unknown. Both imported and locally developed tests are sometimes translated into local languages in order to make the test culturally fair. Translated tests often pose various problems not all of which receive sufficient attention. Some of the problems in translation of tests include vocabulary equivalence, idiomatic equivalence, grammatical- syntactic equivalence conceptual equivalence and experiential equivalence (Sechrest, Fay and Zaidi, 1972). In addition to these problems, reading tests present the researcher with problems of equivalences of writing systems.

1.1. Background

The Zambia Achievement Test was translated from English, in an effort to make the test appropriate for children who do not speak English. Although, up to three years ago, literacy skills were taught in English, teachers in rural areas and government schools in high residential areas usually use the local language that is commonly spoken in the area to teach. Sometimes, in Grades 1, 2 and 3, teachers only use English when they are teaching English as a subject. As a result pupils may go as far as Grade 5 without an

adequate grasp of English (Kashoki, 1990). It is, therefore, only fair to test such pupils' achievement in a language that they are most familiar with. Since the items of the test were translated from English to Nyanja the latter version of the test may not have the same psychometric properties as the original test due to differences between the orthographies of the two languages. Children who have been taught to read in Nyanja only have to learn the sounds of the alphabet with very few exceptions to the rules. In Nyanja, once a child learns the sounds of the alphabet, he or she can read virtually any word. However, Nyanja words can be very long and it is not unusual to find words with three or four consonant clusters in the Nyanja orthography. Words like *nkhokwe* (granary), *mnyontho* (moisture) and *kastwiri* (expert) are not uncommon in Nyanja.

On the other hand, a child who is taught how to read in English will have to learn hundreds of rules that usually define exceptions (Aro, 2004). Some words such as *wagon* are easy to read, as the reader only has to know the sounds of the letters while words like *maneuvers* may not be easy to read. In English the same combination of vowels may be pronounced differently in two words. For instance the vowels *ae* are pronounced differently in the words *dead* and *bead*. The long /i/ sound is spelt differently in the words *machine*, *me*, *fee*, *sea*, *conceive*, *key*, *people*, *subpoena* and *Caesar* (Aro, 2004). This shows that in English a given grapheme can have many different pronunciations and a given phoneme can have different spellings.

Studies show that when a test is translated into another language the psychometric properties may not be the same as those of the original test. Singleton and Vincent (2004) studied the potential reading difficulties faced by Hispanic children in the United States of America and they used translated tests to diagnose dyslexia in different

languages. They found out that the psychometric properties may not be the same when a test has been translated into another language. Singleton and Vincent (2004) point out that the types of tests used for assessing reading in English may not be suitable or applicable for other languages.

1.1.1 Types of Knowledge Relevant for Reading

For a child to be able to read, he or she must be able to have the relevant types of knowledge (Williams, 1998, Trieman, 1993). The following are some of the types of knowledge that are necessary for reading:

- Knowledge of the language of the text or words to be read.
- Knowledge of the world relevant to the cultural context.
- Knowledge of the relevant script and orthography
- Word length - long words are more difficult to read than short words

Knowledge of the Language of the Text to be Read

Reading is difficult because it requires that a person not only be able to speak and understand the language, but he or she must also have the knowledge of parts which the language is made of including words, morphemes, syllables and phonemes (Adams, 1990, Frith, 1980 and Trieman, 1993). With alphabetic writing systems such as English and Cinyanja, awareness of phonemes is especially important. In order to read, a child must be able to segment the words he or she knows into phonemic elements which the

alphabetic elements represent. To do this a child should be aware of the segmentation of the language into units of phonemic size.

Knowledge of the World Relevant to the Cultural Context

Words that are common and frequently used in a particular culture are easier to read than words that are rarely used. A child is more likely to find words that are used in his everyday life easier to read (Trieman, 1993:41; Williams, 1998). Experience with reading also affects reading skills. Children who have access to the print media are more likely to learn to read faster than their counterparts from less privileged homes. Proper names are also easier to read in any given culture. Proponents of the connectionist model of reading argue that the body of words to which a reader has been exposed to is of primary importance in reading (Rayner and Pollatsek,1989).

Knowledge of the Relevant Script and Orthographic Conventions

In order to be able to read, a child should have some knowledge of the writing system of the script to be read.

According to the views of cognitive psychologists reading is carried out through two processes:

- Through the process of word by word memorization and the letters in the words seem to be processed in parallel

- The process of constructing a sound based on the spelling (Rayner and Pollatsek, 1989). For example a child may analyse the word *din* into /d/, /i/ and /n/. This is based on the phonological process of letter- sound translation.

Another dichotomy that was widely accepted by cognitive psychologists of the 1970s and 1980s was the dichotomy between regular and irregular words (Trieman, 1993). Regular words can be easily read while irregular words are difficult to read. *Cat* and *dog* are regular and they can easily be read through the phonological route. *Yacht* and *said* are irregular words and they are more difficult to read.

Word Length

The length of words also affects reading. Long words are generally more difficult to read than short words.

1.1.2 The English Writing system

The English orthography is basically an alphabetic writing system and it represents the spoken language at the level of phonemes. Each phoneme in the spoken form of the word is represented with a letter or a group of letters (a grapheme). However, the English writing system, unlike the ideal alphabet which has one- to- one relations between phonemes and graphemes. In English, written script does not fully represent the phonetic structure of the spoken language (Aro, 2004). The English orthography according to Trieman (1993: 41) is complex in at least four ways:

- It has one-to-many relations from phonemes to graphemes. For example /k/ is symbolized as *k* or *c* and at times as *ck*. The same phoneme /k/ is symbolized as *k* in *kit*, as *c* in *cat* and as *ck* in *mock*.
- For phonemes that have more than one spelling it is not always possible to predict when each spelling occurs. Rules that do exist often have exceptions.
- It has many-to-one relation from phonemes to graphemes. Two or more phonemes are sometimes spelled with the same grapheme. For example, sometimes /θ/ and /ð/ are almost always spelled the same way as in *thin* and *the*.
- 4. Links between graphemes and phonemes are sometimes overridden by morphological considerations

1.1.3 The Cinyanja Orthography

In 1975 the Ministry of Education (MoE) standardized the orthographies of the seven Zambian languages taught as subjects in schools (Icibemba, Kikaonde, Silozi, Lunda, Luvale, Cinyanja and Chitonga) for educational purposes. This was in response to some practical spelling problems in the use and teaching of Zambian languages (MoE, 1975; Kashoki, 1990). Prior to the standardization of the Zambian languages, there were diverse systems of spelling the languages as each Missionary Society introduced an orthography that reflected their background.

Cinyanja has a structure of its own that is similar to other Bantu languages but is very different from the English language and this is usually reflected in the orthography.

Vowels in Chinyanja

There are five short vowels in Cinyanja symbolized as *a*, *e*, *i*, *o* and *u* and five long vowels. The long vowels are symbolized by doubling the vowel. Long vowels are used where the length marks the difference between two words. For example, *bula* (a kind of fruit) is spelled with one *u* and *buula* (to groan with pain).

Consonants

There are a total of 24 consonants in Cinyanja – *b*, *c*, *d*, *f* or *pf*, *g*, *h*, *j*, *k*, *l* or *r*, *m*, *n*, *p*, *s*, *t*, *v* or *bv*, *w*, *y*, *z* or *dz*, *ph*, *th*, *ch*, *kh*, *ŵ*. Despite the fact that *ph*, *th*, *ch*, *dz*, *kh*, *ts*, *ny*, *ng*, *pf* and *bv* are each one phoneme in speech they are each represented by two or three symbols in writing. For example, the labial consonant *ph* in *ipha* (to kill); the velar glottal *ng'* in *ng'ona* (crocodile) and the alveolar *ts* in *tsiku* (day) are all represented by two or more symbols.

In Cinyanja there is a specific rule for writing words with *l* and *r*. The rolled *r* as it exists in the English alphabet. Nearly all words with *l* and *r* have a pure or flapped lateral sound (MoE, 1975: 22). When standardizing the orthography it was decided that the vowels *a*, *o*, and *u* precede *l* and that *r* be preceded by *i* and *e*. This rule applies to all words with the letters *l* and *r* except for words whose stems start with *l*.

For example, *tilemba* (we write) from *ti* and *lemba* does not follow this rule (MoE, 1975). Place names also do not follow this rule.

Double consonants are not unusual in the Cinyanja orthography and they are often found in contractions of words. Long words are also quite common in Cinyanja. This is because prepositions, prefixes and demonstratives for common nouns are joined to the noun. The pronoun *ndi* is joined to the verb. For instance, in the word *ndilikudya*, I am eating.

The information above shows that the two languages are very different from one another. Apart from the differences in the writing system the two languages have totally different language structures.

1.2 Statement of the Problem

Translation of a reading test from a highly irregular orthography, such as English, into a shallow orthography such as Cinyanja is a difficult task as the same levels of difficulty may not be attained on all the test items. In order for the two tests to be equivalent we need to use stimuli of similar difficulty and control for superficial phonological and articulatory language differences.

Is it possible for the translated Cinyanja version of the Zambia Achievement Test-Reading Recognition (ZAT-RR) to have the same psychometric properties as the English version?

1.3 Purpose of the Study

The purpose of the study is twofold:

1. To analyse the basic psychometric properties at item level of the two versions of the test including the reliability of the test.
2. To analyse the content validity and the construct validity of the two versions of the test.

1.4 Specific Objectives

1. To find out if the two tests are comparable
2. To analyse the level of difficulty for each item of the test in the two versions of the test.
3. To compare the mean score for the two groups of children
4. To compare the types of difficulties faced by the two groups in reading.

1.5 Research Questions

1. Are the two versions of the test directly comparable?
2. What is the level of difficulty for each item in the two tests?

3. Is there a difference between the mean score of the children who take the test in English and those who take it in Cinyanja?
4. What type of difficulties does each group of children face?

1.6. Significance of the Study

Reading is something that children are expected to learn in school. It is important for any child to have good reading skills to meet the challenges on learning life. Studies have shown that there is a correlation between reading and other subjects such as mathematics (Nunes, 2000). Children with reading difficulties need to be identified quite early in their school life due to the multiple effect that reading problems can have on other subjects. In order to identify and diagnose reading difficulties it is important to develop a test that specifically identifies a child's reading problem. Tests that have good psychometric properties and can easily used by a researcher are important.

Other studies that have been done in Zambia have concentrated on levels of achievement without looking at what exactly was most difficult for the children and why. This study goes a step further as it analysed each item of the test to find out which items children in the Grades 4, 5 and 6 find most difficult to read. Analysis of the psychometric properties will help establish the actual difficult levels and discrimination indexes of each item. This is necessary for ranking the test so that easier items come first in the test.

Careful development of a test may also help us to identify both gaps and irrelevancies in the materials being taught in schools. If teachers understand the logic behind children's reading they can respond to the errors in the most helpful manner by designing appropriate instruction to overcome the difficulties

1.7 Limitations

The limitation of the study is that the sample was too small for generalizations. As there is a difference in the level of understanding Cinyanja between Lusaka children and children in the other areas where Cinyanja is predominantly spoken such as Chipata the results may not be generalized to all areas where Cinyanja is the lingua franca.

1.8 Operational Definitions

Psychological tests

A psychological test can be defined as devices for quantitative assessment of psychological attributes of an individual. In contrast with the traditional examinations or assessments of a qualitative kind, they are objective and standardized measures of a sample of behaviour (Anastasi 1976). In this report the word test is synonymous with psychological test.

Achievement Test

An achievement test is a type of psychological test that is designed to measure the effects of a programme of instruction (Anastasi, 1976, Cronbach, 1949). Achievement tests measure the effects of relatively standardized set of experiences such as a literacy course of a computer course. Apart from appraising learning, achievement tests can

have both clinical and predictive uses. Achievement tests are often good predictors for making recommendations regarding weaker pupils.

Psychometric properties

Psychometric properties of a test are essentially the various psychological characteristics of a test that make the test behave in a particular manner. According to Coolican (2004:98), “Psychometric properties of a test are various psychological variables of a test ranging from the validity, reliability to the difficulty and discrimination of a test”. In this report the term psychometric properties is restricted to item and test difficulty and discrimination, reliability and construct validity of the test.

Difficulty

Refers to how difficult or how easy an item or test is. It is the percentage of the examinees passing the item or test.

Discrimination

Refers to the effectiveness of an item or a test to discriminate between high and low scorers.

Reliability

Reliability refers to the consistency of a test. Test reliability is the consistency of the scores obtained by the same individual when retested with the same test or with an equivalent form of the test. Before a psychological test is released for general use, it is important to carry out a thorough check of its reliability (Coolican, 2004). Reliability is

dichotomous. It can be external or internal. A test can be both externally and internally reliable or it may be externally reliable but internally inconsistent or vice versa. External reliability is a measure of the stability of test scores in the same individual on different occasions. A reliable test is one that yields the same results consistently in the same individuals. Internal reliability refers to the consistency of a test within itself. High internal consistency generally implies that respondents answer related items in a similar way. If internal consistency is low then the test must be measuring more than one variable. This study only looks at the internal reliability of the test construct validity.

Validity

The validity of a test the ability of a test to measure to measure what it is intended to measure. The validity of a test can be recognized by its face validity or its face value, content validity, criterion validity and construct validity (Pedhazer and Schmelkin, 1991).

Phonological Awareness

Phonological awareness is the awareness of any phonological units of spoken language including syllables, intrasyllabic units, rhymes, onsets and phonemes (Vernon 1971).

Phonemes

A phoneme may be defined as the minimal unit of sound that makes a difference in meaning. In this report all symbols for phonemes are enclosed in slash marks.

1.9. Theoretical Framework

This research is based on the classical test theory. According to this theory it is important for any test that is newly constructed test to be critically analysed to find out whether the test is measuring what it is supposed to measure. The main aspects of this process involve analysing item difficulty and item discrimination of any new test, the reliability of the test, and the validity of the test. The distribution of the test scores must also be examined to see if it produces a normal distribution (Anastasi, 1976 and Cronbach, 1949).

Measuring item difficulty helps the tester to understand what psychological attributes account for the variance in the scores. According to this theory, analysis of items can help us to understand the test as a whole and also understand why a test shows specific levels of reliability and validity (Murphy and Davidshofer, 1991). Every test is constructed on the basis that it measures some psychological construct. For example, it is assumed that a reading recognition test measures reading recognition. Analysis of items in a test is important as it helps us to locate items that do not meet the assumption.

Item Discrimination and Item difficulty

The classical test theory, which is the traditional method of analysing a test, is both qualitative and quantitative. This method uses two main statistics for item analysis – difficulty (or facility) and discrimination. Item difficulty is a measure of the difficulty of an item. Item difficulty is essentially the percentage of persons who answered the item correctly. The easier the item the larger this will be (Anastasi, 1976; 1988 and Aiken, 1994). Item difficulty (p) ranges from 0.00 to 1.00. When item difficulty $p = 1$ then all

pupils taking test got the item right and when $p = 0$ then none of the pupils got the item right.

According to Anastasi (1988) the optimum p value depends on a number of factors:

- The purpose of the test
- The number of response options
- Number of examinees

For a 4- response option test the optimum mean value of item difficulty is 0.69 (Anastasi, 1988). Although several very easy items and several very difficult items are often included in a test, they may actually add very little to the overall effectiveness of the test in differentiating among participants or individuals. As a whole a test should aim to have an overall difficulty of 0.5, although it is acceptable for individual items to have higher or lower difficulty ranging from 0.2 to 0.8

A low item difficulty is ideal for an achievement test that is meant for the selection of a small percentage of examinees whereas a high item difficulty value is ideal for screening out a few poor participants. Therefore for screening pupils with reading difficulty a test with high item difficulty is appropriate (Anastasi, 1988).

Item discrimination (D) is a measure of the effectiveness of an item in discriminating between high and low scores on the test as a whole (Anastasi, 1988, Murphy and Davidshofer, 1991:150). The higher the value of item discrimination the more effective

an item is in discriminating between the high and low scorers. When $D = 1$ then all the examinees in the high scoring group and none in the low scoring group got the item right. When $D = 0$ then the percentages in both groups passing the item are equal. Item discrimination index can have any value between 1.00 and -1.00. When D is negative it means that more examinees in the lower group gave correct responses for the item than examinees in the higher group (Anastasi, 1976, 1988). An item has an acceptable level of discrimination if $D = 0.30$. However, the discrimination index for individual items may range from 0.2 to 1. Item difficulty (p) and item discrimination (D) indexes are not independent of each other and the value of p may vary with the value of D . A value of D smaller than 0.30 is acceptable as p becomes increasingly higher.

In classical test theory, the analysis of multiple - choice items also includes the analysis of the distracters. If the item is well constructed examinees in the upper group will select the correct answer while in the lower group tend to choose the distracters. The magnitude of D indicates the degree of this tendency. When D is negative it indicates that the distracters were selected more frequently by the upper group than the lower group of examinees and such an item needs revising. However, Murphy and Davidshofer (1991) argue that the sign and magnitude of D do not indicate whether all distracters are functioning similarly. Murphy and Davidshofer (1991) and Aiken (1994) suggest that a simple way of determining whether the distracters are working well is to count the number of times that each distracter is selected by the examinees in the upper group and examinees in the lower group. If, on an otherwise satisfactory item, too many examinees in the upper group or few of those in the lower group select a distracter as the right answer, another distracter should be constructed. Ideally all distracters should be

plausible to examinees that do not know the correct answer. All the distracters on a specific item should be selected by approximately the same number of examinees (Murphy and Davidshofer, 1991). According to Kehoe (1995:136), "One should be suspicious about the correctness of any item in which a single distracter is chosen more often than other options including the answer especially if the correlation with the total score is positive".

CHAPTER TWO: LITERATURE REVIEW

This chapter reviews relevant literature on the subject of reading and the problems of translating tests. The literature review will be presented according to the following headings letter knowledge and reading, phonological awareness and reading, problems of translation, effects of orthographies on literacy acquisition, reading studies carried out in Zambia and Zambian languages in the classroom. This chapter also includes a discussion on reliability and validity.

2.1 Letter Knowledge and Reading

Before children can learn how to read, they must be familiar with the alphabetic system of the language. According to Vernon (1971), the ability to discriminate and identify letters is of major importance in reading. Some letters tend to be easier to discriminate than others, for example, very young children have difficulties discriminating between 'd' and 'b'. In learning to read the child must realize that the printed shapes represent the linguistic patterns, which he or she utters and hears. Vernon (1971:27) states that, "The teaching of reading is most effective, if it preceded or accompanied by letter names and sounds."

In order for a child to learn to read he or she must be able to break the words down into their alphabetic constituents, attach appropriate sounds to the printed letters and combine the letter sounds together to form the sound of the word. Studies have shown that children who perform well on letter knowledge also perform well in word recognition (Vernon 1971).

Serpell, Baker and Sonnenschein (2005) asked children to identify letters of the alphabet first in the uppercase then in lower case letters. The letters were presented one at a time on a note card to the child. The child was shown the initial letter of his name and asked to identify it. The remaining letters of the child's name were then shown to the child in a scrambled order for identification. Children's knowledge of specific letters in their names were examined to find out if knowledge of personally significant letters was greater than knowledge of other letters. Serpell, Baker and Sonnenschein (2005) found out that by the time children were finishing kindergarten, they could identify 85% of the letters in their names. It was clear that children were more familiar with letters that were connected with their personal lives. They also found out that those children who performed well on this task also performed well on other reading tasks.

2.2. Phonological Awareness and Reading Skills

Phonological awareness refers to the child's awareness of the sub-lexical segment of speech sounds or the realization that words can be divided into their constituent sound segments. Many studies have shown that the progress children make in learning to read is critically dependent on their phonological skills (Goswami and Bryant, 1990; Liberman et al., 1974 and Ryan, McNamara and Kenny, 1977).

Liberman et al, (1974) investigated the relationship between phonemic segmentation in speech and reading in children. Children were trained to tap on hearing certain sounds. The children listened to each word or sound and tapped out the number of segments with a stick on a table. During training the children observed the correct segmentation of four sets of stimuli. The children were then given 42 items to segment and testing was ended

early if a child got six consecutive right answers. A word recognition task was then given to the children (from the Range Achievement Test) three months later. Liberman et al, (1974) found out that among the poor readers half had not reached criterion in the segmentation task and none of the good readers had failed the segmentation task. Liberman et al (1974) concluded that there was a relationship between phonemic segmentation and early reading acquisition. Similar studies done by other researchers have produced similar results (Ryan Mc Namara and Kenny 1977; Liberman and Shankweiler 1977; Ehri, 1979; Ehri, et al., 2001 and Mann and Wimmer 2002).

Ehri (1979) suggested three ways in which word consciousness might emerge in the beginning reader:

- As a consequence of interacting with print, the beginner suddenly awakens to the fact that meaningful sentences are comprised of word units.
- The beginner may achieve lexical awareness gradually, word by word as he or she decodes and stores the printed forms of the words in memory.
- As a consequence of learning how to spell the reader acquires an orthographic representational system useful for creating visual forms of word sounds which otherwise are transitory.

Ehri (1979) also argued that young children below the age of 7 have difficulty detecting the presence of sound segments in spoken words particularly phonemic units. He also argued that the extent of difficulty depends upon the complexity of the sound-analytic task and the amount of training provided.

2.3 Analysis of Reading Test Results

Many studies have been carried out in reading. The area of reading has received a lot of attention from many researchers as reading is considered to be a very vital part of education for all children who undergo any formal kind of education. According to Cronbach (1949) reading tests have been developed in greater numbers and varieties than other educational tests. Cronbach (1949) also states that reading tests often show many problems in test construction.

Seymour (1986) carried out a study of reading difficulties by children between the ages of 11 and 12 years. The subjects were children from two classes with no reading problems and 22 children with known reading problems. All the children were given an intelligence test (the Weschler Intelligence Scale for Children- Revised) before being assessed on reading. The children were assessed on the following tasks:

- non-word reading
- word reading (vocalisation)
- lexical and semantic tasks
- visual matching tasks.

Seymour (1986) compared the types of mistakes made by the two groups of children and analysed the type of mistakes they made. He then categorised the items on the test into levels of difficulty. He found out that the competent readers differed in a number of ways from the children with reading difficulties. The type of errors made by children who had reading difficulties was different from the children with no reading problems.

This study had the advantage of being able to compare the difficulties faced by children whose reading backgrounds were known.

The other strength of this study is that it classified the levels of difficulties, providing good information on what kind of mistakes children with reading difficulties often make.

Trieman (1993) collected spellings from 43 first-grade children. The children were members of a single teacher's class during two different school years. The school was located in a predominantly white middle-class area in Indianapolis. Spellings of 26 children who attended the teacher's class from one school year and 17 children who attended the class the following year were analysed. The children were between six and seven years old. Their exact ages were not available. Trieman (1993) paired the printed words in each of the children's story with the spoken words in each dictation. She then symbolised the sound of each spoken word. Generally she assigned each letter in the conventional spelling of the word to a phoneme. She found out that children try to spell the words as they sound in isolated form rather than as they sound in connected speech. Although this study gives us a very good picture of the kind of spelling mistakes children are likely to make in the first grade it does not give the reader insight into the processes involved in reading and spelling in higher grades.

2.4 Problems of Translation

In translating material from one language to another there are several problems that one needs to take into consideration.

Some of the problems that Sechrest, Fay and Zaidi (1972) point out are:

- The first problem the cross-cultural researcher has to deal with is the issue of how to introduce the task. The introduction may be very simple in English but very difficult to translate into another language.
- Instructions specific to different tasks or measures being used may not be easy. It is always necessary to translate the instructions in words and when translating short phrases or sentences lack of context may make translation difficult.
- Phrasing questions or other verbal stimuli in ways that are comparable in two languages is difficult to achieve and translation problems may arise even within a culture. One subgroup may not understand the meaning of the material due to differences in dialects.

Translation problems can occur at different levels and there are different kinds of equivalencies and each type of equivalence has different effects and implications. Equivalencies vary in importance and they are most serious when the languages and the cultures are maximally different. One of the most obvious equivalence is in vocabulary, in the words used in the two languages. There may be some terms in one language for which are almost impossible to find in the other language. To find the equivalent form of such a simple word as sad may not be easy in another language. Sechrest, Fay and Zaidi (1972) also observed that dictionary language is often not the language of the people. In order for a translation to be good and equivalent, the translator should be able to know the language used by the prospective respondents.

Another common area that causes problems is conceptual equivalence. Sometimes a word in one language, when translated into another may yield high agreement and yet it may not be that the concepts implied by the two words are identical. A concept that is well understood and frequently used in one culture may not exist at all in another culture. For example, Sechrest, Fay and Zaidi found it difficult to find a concept or even a word, which had the same meaning as the American concept of homosexual in Urdu. In order for translations to be effective from one culture to another they must utilize terms that refer to real things and real experiences which are familiar to both cultures.

Studies have shown that it is difficult to translate psychometric properties from one language to another. A word in English is simply not the same word in terms of difficulty in another language. Therefore, establishing a psychological test with the same psychometric properties across different cultures is very difficult. Equivalence of test items in the source and the target language may not be achieved (Crockett 2004). Translating or adopting tests is not easy and is often not done incorrectly. Tests may, therefore, be valid in one group or culture but may not be suitable for cross-cultural comparison.

Singleton and Vincent (2004) carried out a study on the potential difficulties of reading in children whose first language was not English. They found out that translated tests may not have the same psychometric properties as the original test as the type of tests used for assessing English reading may not be suitable or applicable for other languages. Examples of such tests are tests used in the United States of America, which have been

translated into Hispanic. Singleton and Vincent (2004) further point out the potential difficulties of using translated tests in diagnosing dyslexia in different languages.

2.5 Effects of Orthographies on Literacy Acquisition

Several studies have been carried out to find out the effects of learning to read in different orthographies in recent years. These studies have suggested that basic decoding skills may develop less effectively in English than in some orthographies that are less complicated.

Wimmer and Goswami (1994) carried out a comparative study of reading development in young English and German children. They found out that 7, 8 and 9- year old English children had substantially more difficulties in pseudoword reading tasks than German children.

Landerl (2000) replicated Wimmer and Goswami's study and he found out that young German readers in the first and second grade showed a distinctive advantage in their ability to read pseudowords with a high degree of accuracy. Defior, Martos and Cary (2002) also replicated Wimmer and Goswami's study using Spanish and Portuguese children. They also found out that the Spanish and Portuguese children's performance was relatively similar to the performance of the German children.

Aro and Wimmer (2003) compared reading performance of English children in Grades 1 to 4 with reading performance of German, Dutch, Swedish, French, Spanish and Finnish speaking children at the same grade level. Three different tasks were used; numeral reading, number word reading and pseudoword reading. The results showed that the other orthographies (Finnish, Swedish, German, Dutch, French and Spanish) were read with a high level of accuracy approaching 90 percent by the end of Grade 1.

Seymour, Aro and Erskine (2003) investigated the effects of orthographies in the early phase of reading acquisition through the assessment of letter knowledge, familiar word reading and pseudoword reading in English and 12 other orthographies. The 13 languages were classified into either simple or complex syllabic structure. The languages in the simple syllabic structure class were Finnish, Greek, Italian, Spanish, Portuguese and French. The complex syllabic structure included German, Norwegian, Icelandic, Dutch, Swedish, Danish and English. The results showed that children from the majority of European countries become accurate and fluent in reading before the end of the first year in school. The exceptions were in French, Portuguese, and Danish and particularly in English. The difference appears to be due to differences in syllabic complexity and orthographic depth. The rate of development of reading in English was more than twice as slow as in the shallow orthographies.

Seymour, Aro and Erskine (2003) concluded that syllabic complexity selectively affects decoding where as orthographic depth affects both word reading and non-word reading. It was hypothesised that, the deeper orthographies induced implementation of a dual

(logographic- alphabetic) foundation, which takes more than twice as long to establish as the single foundation required for the learning of a shallow orthography.

Aro et al., (2004) followed-up the acquisition of literacy by 63 Finnish first-graders during the school year. The sample consisted of 34 girls and 29 boys from three different classrooms. The mean age of the children when first assessed was seven years 3 months with a standard deviation of 4.9 months. All the children were fluent in Finnish and one third of them were already accurate readers on school entry. The results of the study showed that development of literacy skills was rapid and at individual level it was characterised by leaps. There was a high correlation between pseudoword reading and word reading and between reading and spelling. The development of accuracy in non-readers was predicted by poor performance in early letter knowledge, phoneme awareness and pseudoword repetition. The results were interpreted as supporting the notion of orthographic difference in reading acquisition.

2.6. Studies Carried out in Zambia

Many studies have been carried out to assess the reading ability of children in Zambia. The review of literature in this study focuses on studies that have assessed reading ability of Zambian children in English and a local language.

Sekeleti (1988) carried out a study to collect factual information relevant to policy decision on whether English should be used as the medium of instruction immediately a child enters school. A sample of 80 grades 5 and 6 children was drawn from each of the

two types of primary school (privileged and underprivileged). He observed speech by pupils and teachers during selected lessons in each of the schools sampled. Specially constructed achievement tests and attitude questionnaires were administered to children and parents. Tests of vernacular language skills and English language skills were used. The tests used included reading tests, a writing test, comprehension tests and comprehension of proverbs.

The English reading test consisted of 25 sentences used in Sharma's (1973) study of grade 3 reading. The Bemba test involved reading short fables of about half a page. These tests were timed and the number of errors and the number of repetitions made were scored on a four-point scale.

The writing test consisted of a sequence of 6 pictures and each subject was asked to describe the events depicted in the picture in English and Bemba. The responses were scored for volume, accuracy and variety. The English comprehension test was based on the Zambia Primary Course Reader IV. It consisted of three multiple-choice questions and 3 short-answer questions. The Bemba test was based on a passage from a Bemba fable. This test had 12 items with 8 of the items being multiple-choice questions. The four remaining items were short-answer-type questions. In the comprehension of proverbs test, the children were required to interpret 10 selected Bemba proverbs from Milimo's (1972) Bantu Wisdom. These were selected after a pilot study in which items, which were very easy and very difficult, were dropped.

The results showed that, “Teachers in the privileged schools used significantly more English when addressing their pupils than in underprivileged schools both in English lessons and in Mathematics Lessons”(Sekeleti, 1988:7).

Privileged school pupils scored significantly higher on the English comprehension test but significantly lower on the Bemba comprehension than underprivileged school pupils. The results also showed a significant difference in favour of pupils from the underprivileged schools in terms of performance in reading and writing in both languages: pupils from underprivileged schools scored significantly higher on the English reading test ($F = 6.58$, $df, 77$, $p < 0.0.1$). This surprising results could have been due to the fact that pupils from the underprivileged schools learnt to read in the local language first, and they could have transferred the skill to reading English. The correlations between the two languages in the three skills that were tested were not significant.

Parents of children from privileged schools scored higher than parents of children from underprivileged schools. Although parents of pupils from privileged schools reported using English more extensively in the home, it was not reflected in their children’s English language skills.

The main flaw of Sekeleti’s (1988) study is that the tests used for testing English and Bemba were not the same. As the psychometric properties of the two tests were not reported there is no way of knowing whether the two versions of the tests that were used

were of comparative difficulty. It is, therefore difficult to compare the performance of the children in the two languages.

The Zambia National Assessment of School Achievement in 5th Grade (ZNASA – 5) was developed by the Ministry of Education as an instrument to be used for the national assessment of grade 5 learning in English and Mathematics. A large sample of Zambian School children (N = 7 844) was assessed using the instrument. Pupils from 400 schools across Zambia were included. This assessment revealed surprising underachievement in both English and Mathematics (Kelly and Kanyika 2000). The mean percentage of scores was 33.2 in English and 34.3 in Mathematics. The results showed that only one in four pupils was achieving at the minimum level (Kelly and Kanyika 2002). The performance was poor in all parts of the Country. Grigorenko et al., (2003) used the G5NA - 5 in the Eastern Province of Zambia to assess about 1200 children. The level of performance in this sample was comparable to that reported by Kelly and Kanyika (2000) in Mathematics (33.8%) but significantly lower in English (12.6%).

Grigorenko et al., (2003) built on these findings and constructed two parallel versions of a written vocabulary test, one in English and one in Chinyanja. This test was administered to 387 middle school pupil in Zambian public schools. The results showed the levels of performance of children in both English and Chinyanja was low. The mean for the performance in English was 9.8 with a standard deviation of 4.6 and for Chinyanja the mean was 10.7 (SD = 8.0). Grigorenko et al., (2003) concluded that the knowledge of grade appropriate vocabulary was low in both English and Chinyanja. They also concluded that the results might have been confounded by poor reading

scores of the Zambian children as the vocabulary words were presented in a written multiple-choice format.

Williams (1998) carried out two studies in 1992 and 1994 to investigate the reading

proficiency of pupils in English and a local language (Chinyanja in Zambia and Chicewa in Malawi) in *Zambian and Malawian Primary Schools*. In the 1992 study he investigated reading proficiency at year 3, 4 and 6. In this study particular attention was paid to the effect of location, urban or rural, and gender differences. The study also investigated the teaching of reading in schools in the two Countries. Thirty-two pupils were tested, 480 Malawian pupils and 452 Zambian pupils. The mean ages for Malawi and Zambia were as follows:

<u>Year</u>	<u>Malawi</u>	<u>Zambia</u>
3	12.3	11.4
4	13.4	12.5
6	14.10	13.9

Classroom observations were carried out and group reading tests were administered in English and the local language at year 3, 4 and 6 in five schools in each Country. None of the children in the sample spoke English at home. The instruments used for assessing the pupils were the Modified Close Test for Chicewa and Chinyanja.

In the Modified Close Test (in English) the reader was given a passage to read with four to six gaps and he or she was asked to fill in the gaps with words provided in a box.

This word-find test was divided into three parts of 20:

Items 1 – 20 Year 3

Items 21 – 40 Year 4

Items 41 – 60 Year 6

All 60 items were given to all the pupils who took part in this study. The reliability for this test was 0.86 (Kuder Richardson – 21).

The Local Language Modified Close Test was a Chicewa and Chinyanja version of the word find test. Both versions were adaptations of the English test. The reading tests were given to whole classes where the class was fewer than 34 and in larger classes a representative sample was selected. Results showed that pupils do improve with time. The study also showed that most Zambian children had problems in reading their local language. Malawian children scored significantly higher in reading in the local language. The results also showed that pupils who scored high in English also scored high in Chinyanja. Correlations between English and the local language results were positive for Malawi. The reliability for the local language version of the test was not reported.

Williams (1994) carried out a similar study in 1994 to look at comprehension strategies of low and high groups, the difficulties reported by the pupils in reading and ways of overcoming them. The sample was drawn from grade 5 primary school pupils from

Malawi and Zambia. There were 270 participants from Malawi and 277 from Zambia. The data collected included a date of birth, sex and home language. The group reading test was administered first, followed by a break of 15 to 30 minutes then the local language test was administered. The reliability for the English test was 0.84 Kuder Richardson for Malawi and 0.95 for Zambia. For the local languages version of the test the reliability index was 0.79 (Kuder Richardson – 21) for Malawi and 0.75 for Zambia.

Some of reasons the examinees in Malawi gave for preferring to read in Chicewa was that they perceived it to be easier to read than English. The majority of the Zambian pupils said they preferred to read in English and yet they had difficulties reading English. The results showed that low scorers had considerable difficulties with vocabulary.

William's studies were well thought out and also brought out some of the problems that pupils in the primary schools have in the area of reading. However, it did not illustrate why English was more difficult to read than Chicewa and Cinyanja.

2.7 Zambian Languages in the Classroom

Ohannessian (1976) carried out a study in Zambian schools in or near Chipata, Kabwe, Kalomo, Kitwe Lusaka, Livingstone, Mufulira and Ndola to observe the teaching of Zambian Languages in Primary schools in 1970. The observations were carried out in 254 classes in 106 schools in which 5,592 boys and 4167 girls were observed (a total of 9,759 pupils). The schools were classified into large town, medium sized town small towns and country areas. Observations were made of Zambian language lessons in

grade 1 to 7. The observers recorded the language being taught in the classroom and the regional or approved language for the area. The following observations were made:

- In some cases a language other than the seven official languages was taught, for example lessons in Tumbuka were observed in the Eastern Province. In one case a teacher was reported as speaking Lenje throughout the lesson, which was recorded as Tonga on the timetable.
- Sometimes lessons were in a language that was not the approved language or a mixture of other languages. For example one teacher is reported to have used Bemba and Nyanja as a medium of instructions in a Tonga lesson
- There were some reports of Zambian languages not being taught at all. One teacher on the Copperbelt reported that she taught arithmetic during the Zambian Language lesson, as she did not know the language to be taught.
- In spite of not knowing the language many teachers make an effort and learn the language so that they can carry out their work.

Although the actual lessons reported in languages other than those officially recognised was very small, Ohannessian (1976) believed that a considerable amount of such teaching was taking place.

In each class observed all the children were asked to give their mother tongue. A certain amount of reluctance was observed among the children to disclose their mother tongue in mixed urban areas. The results showed that the number of different languages in a classroom varied with locality. Seventeen languages were claimed in one class in

Mindolo in Kitwe. The number of languages was more in the Copperbelt and towns along the line of rail. In the rural areas, occasional classes were unilingual. These results showed that only 82 of the 254 classes had 80% of the pupils who spoke the same language at home, and most of these schools were in rural areas.

Ohannessian (1976:312) concluded that,

‘The problems related to language background are in reality more complex than has been evident so far. Not only is there a multiplicity of language backgrounds, it is not at all easy to determine to what degree some of these languages can be regarded as distinct languages or other dialects of the language spoken by those who claim to speak them as their own mother tongues.’

This research illustrates very well how language backgrounds of both the teachers and the pupils can be an obstacle in the teaching and learning of local languages in Zambia.

McAdams (1976) also pointed out that one of the major difficulties of teaching Zambian languages in urban schools was the number of languages spoken in one classroom. A survey done in 1966 in Lusaka showed that only 49% of pupils spoke Nyanja as their home language and only two of the 24 teachers interviewed were Nyanja speakers (McAdams 1976:336).

Serpell (1970) carried out a field experiment on the comprehension of Chi-Nyanja by Lusaka School children. A total of 1,500 grade 3 and 6 pupils were interviewed to find out their parent language or mother tongue. A sub-sample of 250 children from single language families was tested individually on Nyanja comprehension. A number of tests were used: action, identifying objects and pictures and grammar with nonsense words. The sample was drawn from three large schools adjacent to Libala and three in Matero. The age range of the children was 8-11 for grade 3 and 12 to 15 for grade 6. Results showed that native speakers of Nyanja were superior to all other groups, Bemba speakers experienced less difficulties than Tonga and Lozi speakers. The length of stay in Lusaka also affected performance on the tests. Children from Matero performed better than the children from Libala.

This study shows that although children who are not native speakers of Cinyanja pick the language from their environment, their understanding of the language is not very good.

Reliability

Items in a test are usually correlated and there is often a correlation between test performance and the items. In order for a test to be of any practical use it must have high internal consistency, that is, items that measure the same construct must hang together (Anastasi, 1976 and Coolican, 2004). Inter-item correlation is a general indicator of test quality. In classical test theory the reliability of a test is usually tested by looking at the item – total correlations. Items that have a correlation lower than 0.15 with the total score are not good items as they lower the internal consistency of the test. Items that have negative correlations with the total score are also bad, as they do not

measure the same thing that the total score is measuring. In addition to analysing item-total correlations, item- by- item, the overall internal consistency has to be computed using the Cronbach Alpha coefficient. Cronbach's Alpha depends, largely, on how the examinees vary on individual items. If the examinees vary a lot on individual items relative to how much they vary on the overall test. Coefficient alpha of a test reflects the extent to which the test would yield the same ranking of examinees if re-administered with no effect from the first administration of the test (Kehoe, 1995 and Coolican, 2004). A test with an alpha value of 0.75 to 1.00 has acceptable reliability.

Validity

The validity of a test the ability of a test to measure to measure what it is intended to measure. The validity of a test can be recognized by its face validity or its face value, content validity, criterion validity and construct validity (Pedhazer and Schmelkin, 1991). The face validity of a test is whether a test is obviously measuring what it is intended to measure. For example, a test on reading that has most of its items on arithmetic has no face validity. Content validity is the ability of a test to sample behaviours that are representative of the attributes being measured.

Construct validity is the ability of a test to measure the psychological constructs that it is intended to measure. Before a test can be used with confidence, in addition to establishing that it has high reliability and an appropriate spread of difficulty, some objective evidence is needed as to its usefulness (Cronbach, 1949). Construct validation helps to determine

whether test scores provide a good measure of a specific construct. Construct validation attempts to understand what is being measured by examining the relationship between constructs. It involves understanding why items are related by examining the underlying concepts. Construct validation implies testing out the theory underlying the test and it helps to determine the psychological meaning of a test. For example, reading recognition is a measurable construct and there should be factorial consistency in a test of reading recognition. Tests are designed to measure psychological constructs. According to Pedhazier and Schmekin (1991) psychological constructs have two essential properties:

- Constructs are abstract summaries of some regularity and they are related to concrete observable entities or events.
- Psychological constructs are always related, directly or indirectly to behaviours or experience.

The items in the test should be consistent with the structure and a set of indicators should reflect the same construct. Indicators of a given construct must be correlated and they must have a high loading on the construct being measured. It is important to analyse the construct validity of any new test so that we are confident that the results of the test are due to the construct and not due to error. Negative correlations between the tasks in a test that is measuring one construct shows that there is a weakness in the construct. Any task that correlates negatively with the intended construct shows that the construct is weak. In short tests must be consistent with the theoretical structure.

Summary of the Reviewed Literature

There was a lack of literature on the instruments used in Zambia to assess reading. Literature on the psychometric properties of translated tests was also lacking. However, other sources suggest that reading acquisition is dependent on letter knowledge, phonological awareness, and the writing system. In spite of the fact that most of the literature comes from settings that are different from Zambia it is still enlightening on the subjects of literacy acquisition, and problems of translating tests. This is because there are very few differences in the way children learn how to read across cultures. Some of the types of knowledge relevant for reading that research has identified include letter knowledge, phonological awareness, knowledge of the language of the text to be read and the writing system.

Studies suggest that letter knowledge is very important in learning to read. Children must be able to identify the letters of the alphabet before they are able to read. They must realise that the printed symbols represent linguistic patterns of speech. In alphabetic writing systems, such as English and Cinyanja, letter knowledge is a crucial aspect of learning to read (Trieman, 1993).

A number of the sources reviewed place a lot of emphasis on the ability of a child to *segment speech sounds as a prerequisite for reading acquisition (Goswami and Bryant,*

1990). Studies have also shown that practice on reading tasks has an effect on reading acquisition. The more a child practices on tasks of reading the faster the child will learn how to read. Children who have access to reading materials and have an opportunity to practice reading have a better chance of learning to read.

With respect to writing systems, research shows that when a child learns to read in an orthography that is regular that child will be able to read in the first year of school. On the other hand children who learn how to read in a complicated writing system such as English may take up to four years to read.

The studies that were reviewed also suggest that translation of reading tests is not an easy task. When a test is translated from one language to another, it may not have the same psychometric properties as the original test. Attaining equivalencies in the vocabulary, syntax and concepts is difficult.

Studies of the language backgrounds of pupils in Zambian schools suggest that although children may live in an area where Cinyanja is the lingua franca, such as Lusaka, not all the children speak the language at home. These children may, therefore, have problems understanding Cinyanja.

In conclusion it could be said that translation of a reading test from one language to another can be a difficult task. The translated test may have very good face validity but the ability of the two tests to discriminate between children who have problems in reading and those who are good readers may not be the same when the writing systems of the two languages are different.

CHAPTER THREE: METHODOLOGY

This chapter dwells on the methods that were used to collect the data for the study. It gives an account of the sample used, instruments, data collection procedure and data analysis.

3.1 Sample

The sample consisted of 240 school children, 121 in the English group and 119 in the Chinyanja group. The 240 pupils were randomly selected from two schools in Lusaka, Olympia Basic School and Ng'ombe Basic School and assigned to either the Chinyanja or English group depending on which language they felt most comfortable with. A short screening instrument was used to determine which condition to assign the children to. The distribution of the sample was as follows:

GRADE	Number of pupils tested in English		Number of pupils tested in Cinyanja	
	OLYMPIA	NG'OMBE	OLYMPIA	NG'OMBE
4	32	9	21	19
5	33	6	10	30
6	28	13	19	20

The age ranges of the children were between 8 and 13 for grade 4, 9 and 18 for grade 5 and 10 and 15 for grade 6. . All the pupils in the sample went through the English immersion programme.

3.3. Instruments

The assessment tool that was used was the Zambia Achievement Test Reading Recognition (Cinyanja and English versions). The test was constructed by Dr Tina Newman and Dr Jonna Kwaistkowski and has not yet been published. This test consists of sixty (60) items ranging from letter recognition to reading single words. Items 1 to 40 are multiple-choice questions and a child has to choose an answer from four options. Items 41 to 60 consist of four words on each card and a child has to read all of the four words on the card aloud. This test consists of five tasks: letter matching, sound matching, letter sound matching and word recognition tasks. The test is an individual test that is given to one pupil at a time. A short questionnaire with six questions was used to select the children in each group.

3.4 Procedure

The data were collected from two schools Ng'ombe and Olympia basic schools in November 2004. The data were collected over a period of two weeks. First an atmosphere of friendliness between the tester and the pupil to be tested to make the pupil comfortable and to that he or she ensure co-operation was created. The demographic information including age, type of house, family size, availability of books in the house and which language the child was most comfortable with. The test was introduced to the examinee by explaining that they would be given a reading test and some of the items would be easy and others would be hard. The examinee was asked to try and answer all the questions even if they were hard. The examinee was then taken through the training

items to ensure that he or she understood the instructions. Answers to the questions were recorded. The responses were scored 0 or 1.

3.5 Data Analysis

Quantitative and qualitative methods of analyses were used to analyse the data. Group means were calculated and a t-test was used to compare the means for each grade level using SPSS. Factor analysis, ANOVA and internal consistency of the two versions were all computed using SPSS.

Item difficult and discrimination indexes were also calculated using the following formulas:

$$\text{Item Difficulty index, } p = \frac{U_p + L_p}{U + L}$$

Where p stands for item difficulty,

U_p is the number of test takers in the upper group who passed the item,

L_p is the number of test takers in the lower group who passed the item,

U and L designate the total numbers in the upper and lower groups respectively.

$$\text{Discrimination index, } D = \frac{U_p - L_p}{U}$$

U

Where D is the discrimination index of the item,

U_p and L_p are the numbers of test takers in the upper and lower groups respectively who passed the item,

U and L represent the numbers of test takers in the upper and lower groups respectively.

The distracters were also examined to find out whether they were working effectively.

A second level of analysis was to look at the items that proved to be particularly difficult in both versions of the test. These items were selected on the basis of the discrepancy between the high performers and the low performers. Both quantitative and the qualitative analyses were carried out on these items.

CHAPTER FOUR: RESULTS

This chapter focuses on the results obtained in the study. The results reported include item difficulty and discrimination indexes, reliability coefficients, factor loadings, means, standard deviations and t-values.

4.1. Item Difficulty and Item Discrimination

The mean item difficulty (p) for the Chinyanja test was 0.62 and the discrimination index (D) was 0.62. For the English version the difficulty and discrimination indexes were 0.58 and 0.48 respectively. The mean difficulty and discrimination indexes for Letter Identification, phoneme identification and word recognition were as follows:

	Letter Identification		Phoneme Identification		Word Recognition	
	English	Nyanja	English	Nyanja	English	Nyanja
Mean p	0.94	0.94	0.80	0.66	0.45	0.53
Mean D	0.05	0.06	0.21	0.26	0.66	0.85

The phoneme identification tests were more difficult for the Chinyanja group whereas the word recognition items were more difficult for the English group.

Items 82, 83, 85, 87, 88, 89, 90, 92, 93, 94, 95, 96, 97, and 100 were more difficult for the Chinyanja group of children.

Item 95 was very difficult for both poor readers and good readers in this group ($p=0.03$).

The choice of options was as follows:

Option	Frequency	Percent
A	14	11.8
B	42	35
C	35	29.4
D	28	23.5

The correct answer was option C but 35% of the respondents selected B as the right answer. Nine of the pupils in the upper group chose this option as the correct answer and 11 were from the lower group. The rest were from the medium group. Option B was a very attractive choice and it was a bad distracter as it attracted both those in the high achieving group and in low achieving group. More pupils chose this option although it was a very unlikely option. The option B is a picture of a rock but it does not look like a rock at all. Items 115 to 121(see Appendix C) were more difficult for the English group than the Chinyanja group. Words like anchor and silhouette were very difficult for the English group.

There were a few items with negative discrimination indexes in both tests. Items 72, 74 and 75 had negative discrimination indexes in the Chinyanja version of the test and items 70 and 75 had negative discrimination indexes in the English version. Distracter analysis in the Chinyanja results showed that 98.8% of the pupils chose the right answer for items 72 and 75 and 99.2% gave the correct response for item 75. For item 75, only one pupil gave a wrong response and that pupil was one of those who performed well on

the test. For the English test, 98.8% got items 70 and 75 correct. One of the two pupils who chose a wrong response was from the upper group.

4.2. Reliability Results

The reliability coefficients for the test were very similar. The alpha Cronbach values for the two versions of the test were 0.98 and 0.97 for Cinyanja and English respectively. The items on both tests were quite highly correlated.

A few items in both versions of the test had negative correlations with the total score. In the Cinyanja test 69, 71, 72, 74, 75, 81, and 86, had negative correlations with the total score. In the English version, items 64, 69, 70 and 75 had negative correlations with the total score. These items also correlated poorly with the total score ($r = -0.03$ to -0.17). None of these items had a significant correlation with the total score and apart from one item (item 81) in the Chinyanja version had a significant correlation with the total score ($r = 0.17$, $p = 0.05$). Although these items are exactly the same in the two tests 6 pupils in the Chinyanja group chose a distracter rather than the right answer. The distracters were equally attractive with two pupils choosing each distracter. This item appears to function differently in the two versions of the test.

These items appear not to be testing the same skills that the whole test is supposed to test. Items 69 and 75 had negative correlations in both versions of the test. Further analysis of the items correlating negatively with the total score in the English version showed that most of these items came quite early in the test and only one or two of the pupils chose a wrong distracter.

In general, items 62 to 101 correlated poorly with the total score in both tests. Correlations with the total score for these items were between 0.03 and 0.39. Item- total correlations of items 102 to 121 were between 0.45 and 0.81.

The English version of the test showed significant correlations between all the tasks and the total score. The lowest correlation was between letter identification and phoneme identification ($r = 0.189, p < 0.05$). The correlation between letter identification and word identification was more significant ($r = 0.246, p < 0.01$). Letter identification also correlated significantly with the total score ($r = 0.31, p < 0.01$). There were significant correlations between phoneme identification and word identification and the total score ($r = 0.41, p < 0.01$ and $r = 0.55, p < 0.01$ respectively). The highest correlation was between word identification and the total score ($r = 0.99, p = 0.00$).

The Chinyanja version showed a different pattern of correlations between the tasks in the test. Unlike in the English version, there was no significant correlation between the letter identification and phoneme identification. However, letter identification correlated significantly with word identification ($r = 0.199, p = 0.015$). There was a significant correlation between phoneme identification and word identification ($r = 0.408, p = 0.000$). The total score correlated with all the tasks. The highest correlation was between word identification and the total score ($r = 0.99, p = 0.000$).

4.3 Factor Analysis Results

Factor analysis results showed that both versions of the test were basically testing one component. The loadings for both tests were high on word recognition. The loadings on word identification for Chinyanja were higher than for English.

Component Matrices for the two tests

	English	Chinyanja
	Component 1	Component 1
Letter matching	0.594	0.497
Phoneme identification	0.764	0.755
Word identification	0.798	0.812

4.4 Distracter Analysis Results

An analysis of the options was carried out to find out which distracters were more attractive to both the examinees in the lower group and those in the high scoring group. Ideally the scores should be evenly distributed for all the distracters. Analysis of distracters showed that for items 62 to 76 some distracters were not chosen by any of the examinees. In the Chinyanja test the following items had bad distracters:

Item No.	Distracter	Frequency	percent
88	D	44	37.1
91	A	41	34.5
98	A	25	21.0
66	A	11	9.2
78	A	15	12.6
80	D	14	11.8

D for item 88 and A for item 91 were exceptionally bad distracters as both those from the high and low scoring groups selected it as the right answer.

In the English version the following were bad distracters:

Item No.	Distracter	Frequency	Percent
78	A	21	17.4
80	C	10	8.3
85	A	19	15.7
89	D	22	18.2
98	D	32	26.4

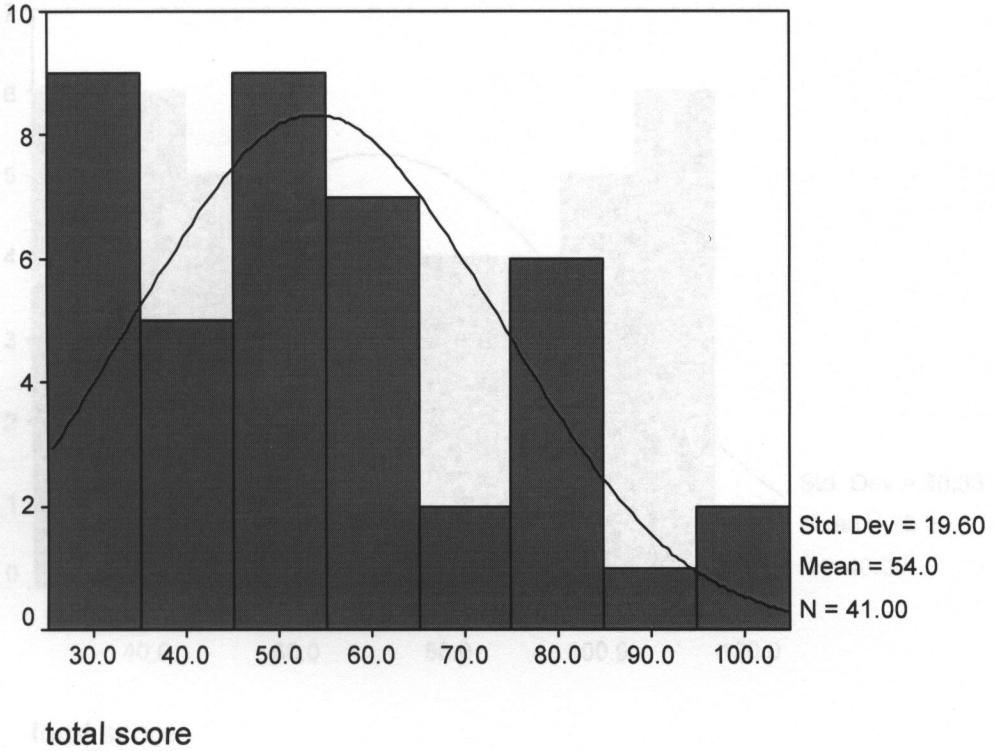
Item 98 was a very bad distracter in this test

Means, Standard Deviations and Distribution of Scores

The means and standard deviations for each grade in each group were computed using SPSS. Graphs were also drawn in order to find out if the test produced a normal curve.

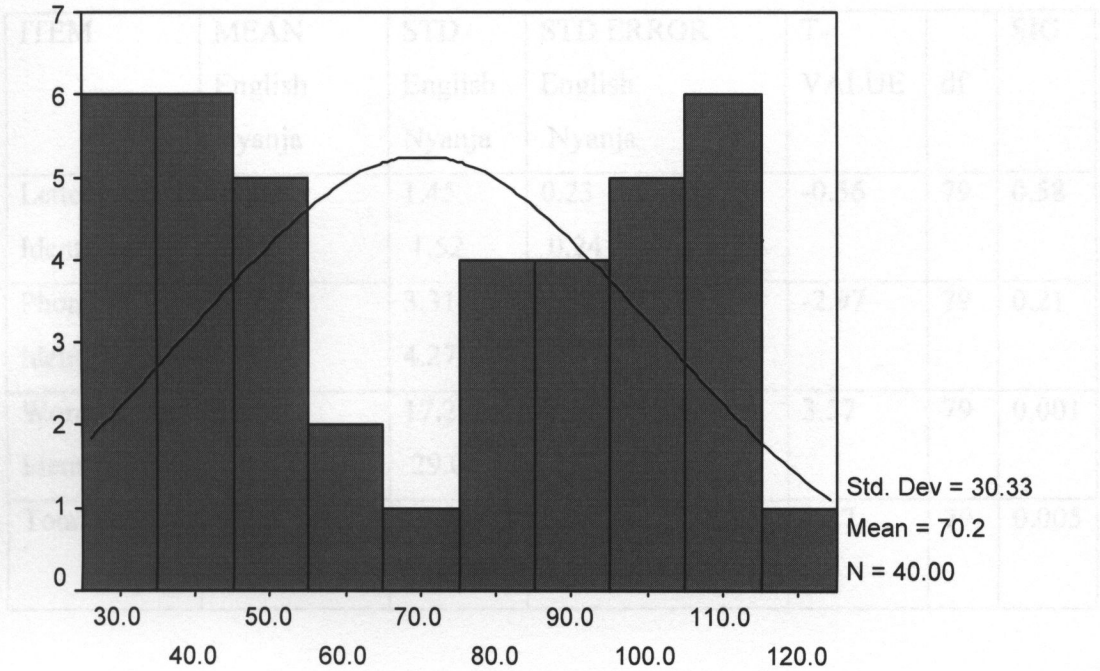
The distributions of the total score for each grade was analysed and compared.

Figure I: Grade 4 English Total Score



The graph distribution of grades 4 total scores for the pupils who took the test in English. The graph is negatively skewed with most of the scores falling between 25 and 60. The mean was 54 and the standard deviation was 19.60. with very few scores around the mean. About half of the pupils performed well and half were below the mean. The large piling of scores on the left is an indication that many examinees in this group found the test difficult.

Figure II: Grade 4 Nyanja Total Score



total score

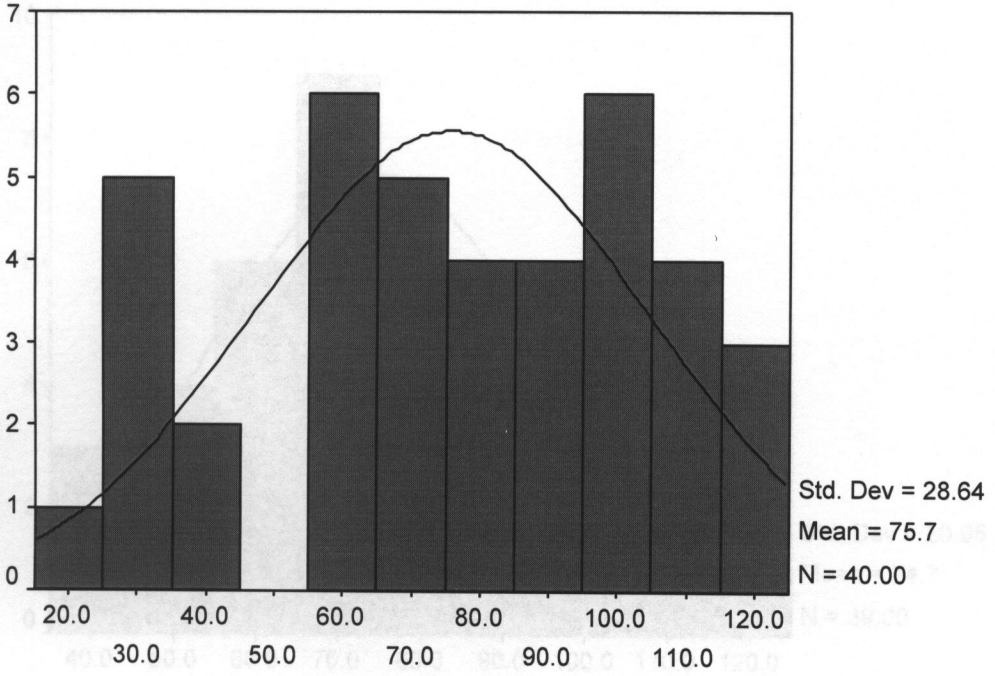
The above graph shows the distribution of scores for the Grade 4 pupils who took the test in Chinyanja. The graph showed a bimodal distribution with very few scores around the mean. About half of the pupils performed well and half were below the mean. The large piling of scores on the left is an indication that many examinees in this group found the test difficult.

Table 1:Table of Sigures showing the Grade 4 Means Standard Deviations Standard Error of Mean t-test Values and Significance

ITEM	MEAN	STD	STD ERROR	T-	df	SIG
	English Nyanja	English Nyanja	English Nyanja	VALUE		
Letter Identification	28.63 28.45	1.45 1.52	0.23 0.24	-0.56	79	0.58
Phoneme Identification	15.51 13	3.31 4.27	0.52 0.37	-2.97	79	0.21
Word Identification	19.83 38	17.29 29.05	2.70 4.60	3.57	79	0.001
Total Score	53 70.2	19.60 30.33	3.06 4.70	2.87	79	0.005

There was no significant difference between the means of the two groups in the Letter Identification and Phoneme Identification ($t(79) = 0.56, p = 0.58$ and $t(79) = 2.97, p = 0.21$ respectively). But the mean for Word Identification and the total scores were significantly different in the two groups ($t(79) = 3.57, p = 0.001$ for word identification and $t(79) = 2.87, p = 0.005$ for the total score). The pupils who took the test in English scored higher on the Phoneme Identification task than the Chinyanja group, whereas the mean score on word Identification was lower for the English group.

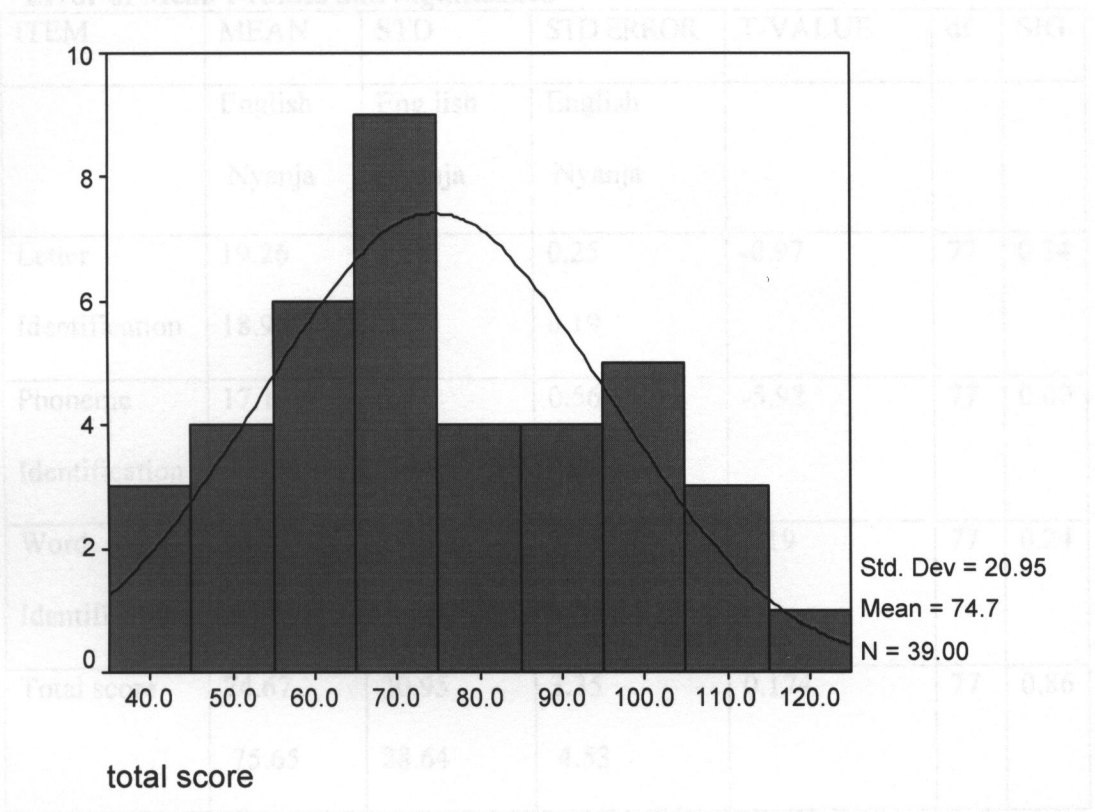
Figure III: Grade 5 Nyanja Total Score



total score

The above graph showed the distribution of total scores for the Grade 5 Chinyanja group. The graph was skewed to the right with most scores falling between 55 and 100. Eight (8) pupils fell 1.5 standard deviations below the mean. There were no scores between 45 and 55. The graph showed a very distinct demarcation between those who performed relatively well and those in the low scoring group.

Figure IV: Grade 5 English Total Score



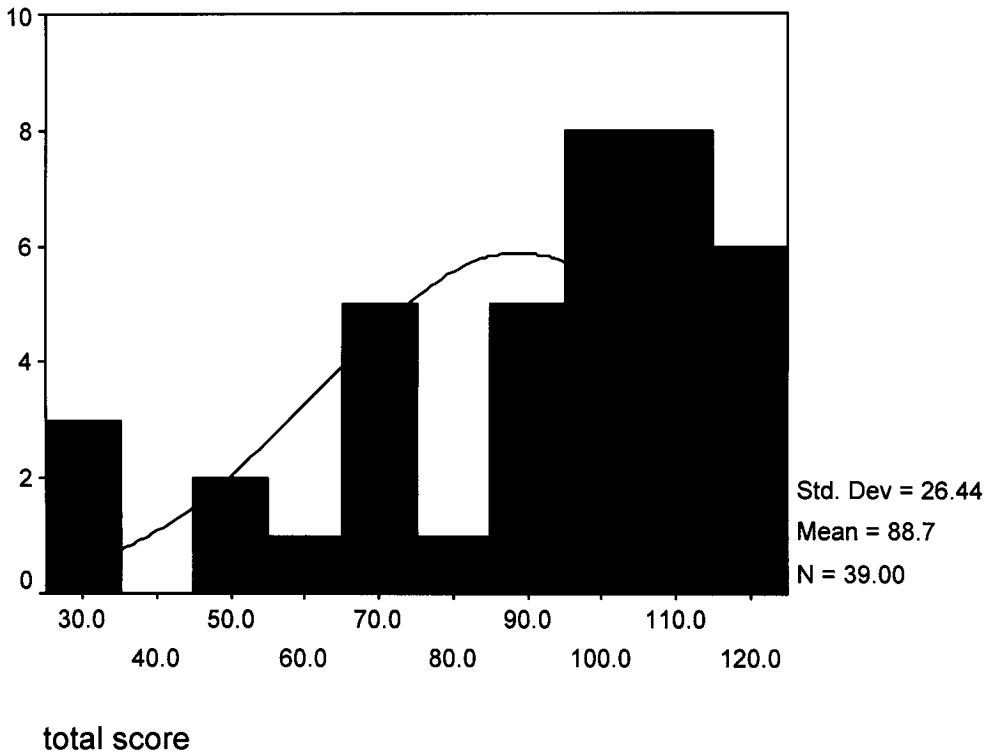
The above graph shows the distribution of the total scores for the Grade 5 pupils who took the test in English. The distribution of scores was slightly skewed to the left. More than half of the scores fell below the mean. The mean score was 74.7 (SD = 20.95, N = 39). Very few pupils had very high scores in this group.

Table 2:Table of Figures Showing Grade 5 means Standard Deviation Standard Error of Mean t-values and Significance

ITEM	MEAN	STD	STD ERROR	T-VALUE	df	SIG
	English	Eng lish	English			
	Nyanja	Nyanja	Nyanja			
Letter	19.26	1.58	0.25	-0.97	77	0.34
Identification	18.95	1.22	0.19			
Phoneme	17.13	3.51	0.56	-5.92	77	0.00
Identification	12.15	3.95	0.62			
Word	38	19.67	3.15	1.19	77	0.24
Identification	44.55	26.50	4.19			
Total score	74.67	20.95	3.35	0.174	77	0.86
	75.65	28.64	4.53			

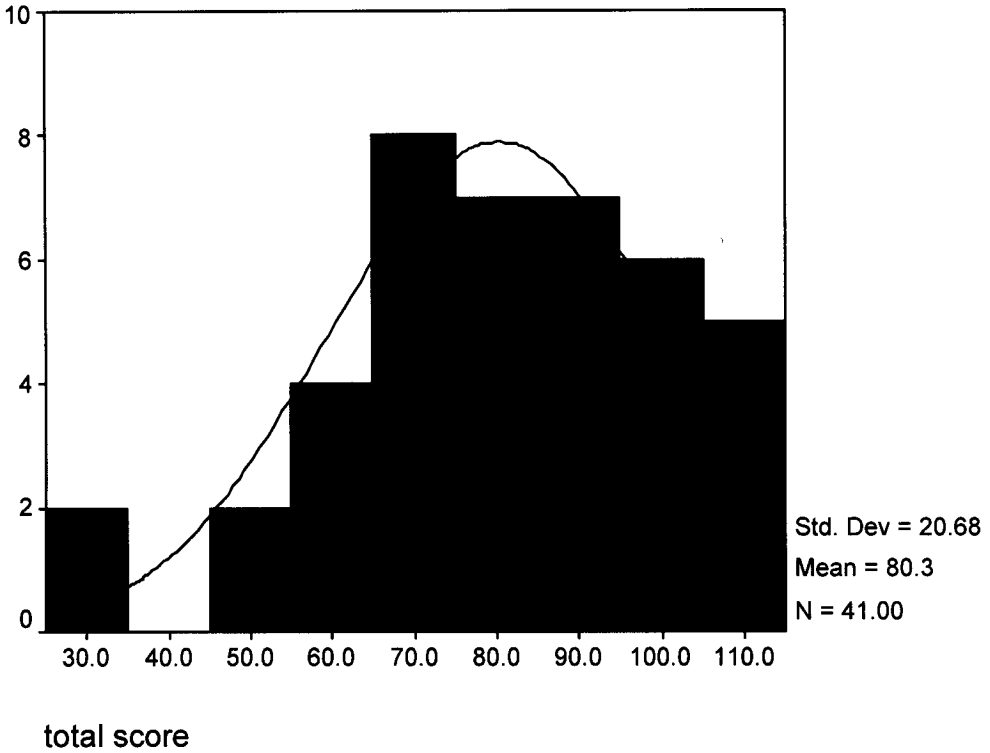
There was no significant difference in the performance of the two groups on Letter Identification and Word Identification tasks ($t(77) = 0.97, p = 0.34$ and $t(77) = 1.19, p = 0.24$ respectively). But there was a significant difference in the performance of the two groups on the Phoneme Identification task ($t(77) = 5.92, p = 0.00$) in favour of the English group.

Figure V: Grade 6 Nyanja Total Score



The above graph shows the distribution of the total scores for the Grade 6 pupils who took the test in Chinyanja. The distribution of the scores was positively skewed. The mean score was 88.7 (SD= 26.44, N = 39). The graph showed that for grade 6, most of the pupils fell above the mean. The graph shows that some of the high scorers got almost all the test items right.

Figure VI: Grade 6 English Total Score



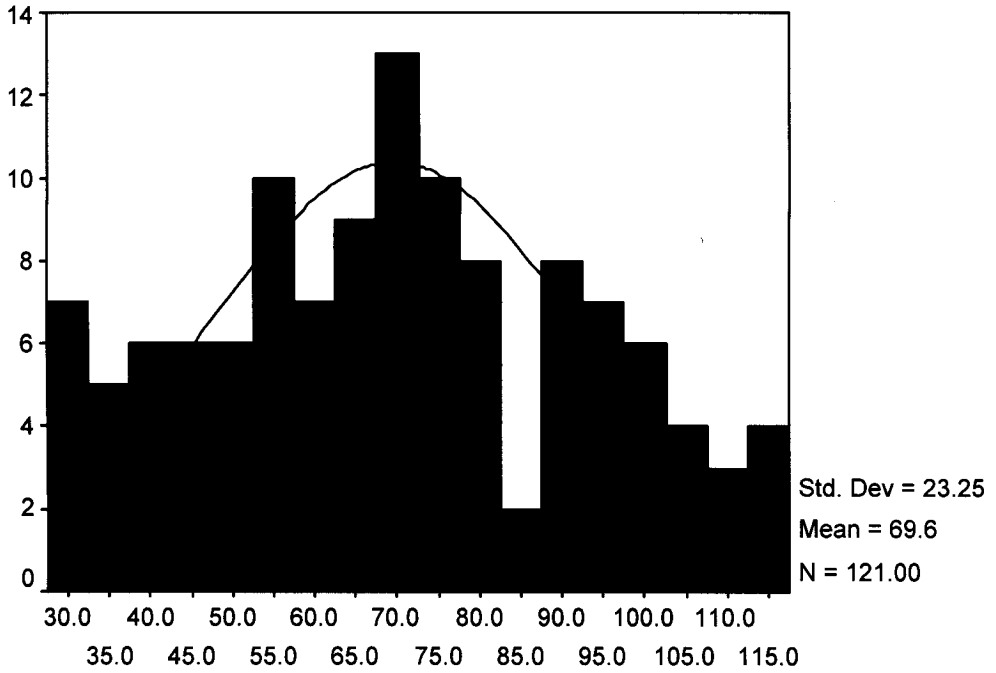
The above graph shows the distribution of Grade 6 total scores for the English group. The distribution of the scores shows the scores are skewed to the right with about two pupils on the extreme right. The mean was 80.3 with a standard deviation of 20.68. There were no scores between 35 and 45. Two of the examinees got below 35 and the highest got 115. The piling of scores on the right was an indicator of the relative facility of the test for this group of pupils.

Table 3: Table of Figures Showing Grade 6 Means Standard Deviation Standard Error of Mean t-values and Significance

ITEM	MEAN	STD	STD ERROR	T-VALUE	df	SIG
	English Nyanja	English Nyanja	English Nyanja			
Letter Identification	19.26 19.25	0.98 1.98	0.15 0.18	-0.05	78	0.96
Phoneme Identification	17.02 14.64	4.61 3.75	0.72 0.60	-2.53	78	0.07
Word Recognition	44 54.80	18.11 24.07	2.83 3.85	2.27	78	0.03
Total score	80.32 88.69	20.60 26.90	3.23 4.23	1.58	78	0.12

There was no significant difference in the two groups of pupils on letter identification and phoneme identification at Grade 6 level although the mean score for the English condition was higher. The pupils in the Chinyanja group performed significantly better than the pupils in the English condition ($t(78) = 2.27, p = 0.03$).

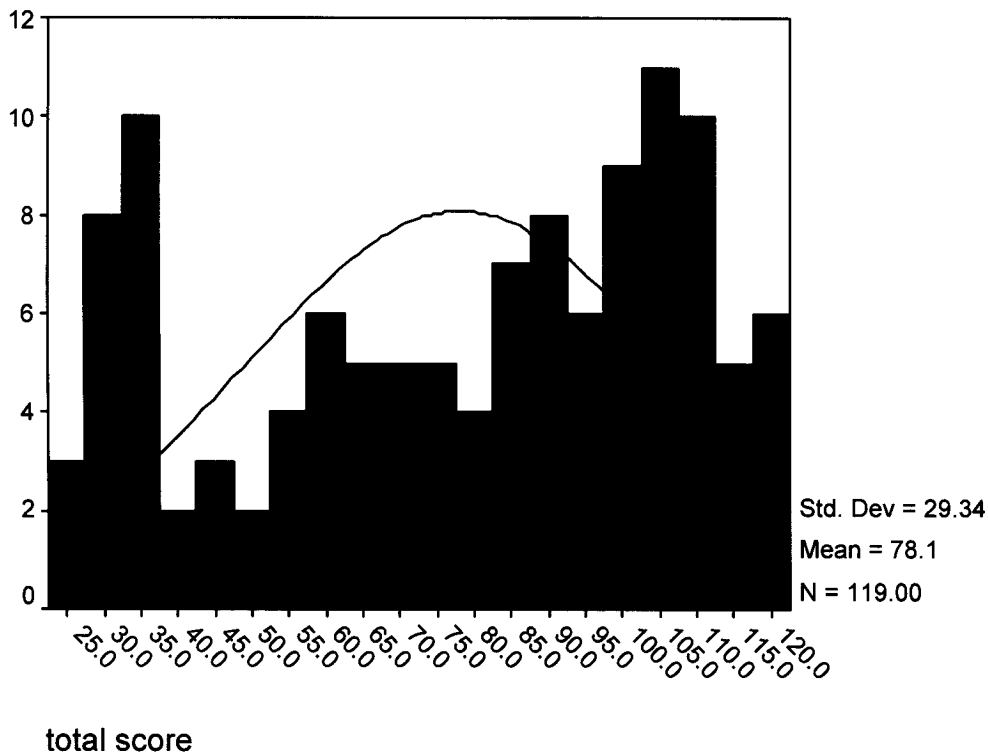
Figure VII: English Grades 4, 5, and 6 Total Score



total score

The above graph shows the overall performance for the pupils who took the test in English. The distribution of scores in this sample of pupils was relatively normal. The mean was 69.6 and the standard deviation was 23.25 (N= 121).

Figure VIII: Nyanja Grades 4, 5 and 6 Total Score



The above graph shows the distribution of scores in the Chinyanja group. The mean score was 78.1 (N= 119, SD =29.34). There were relatively few scores around the mean. The graph shows that there is a group of pupils who were more than one and a half standard deviations below the mean.

Table 5; Table of Figures for Analysis of Variance

	SOURCE	Sum of squares	df	Mean square	F	Significance
Letter Identification	Grade	22.95	2	11.48	6.57	0.002
	Language	1.67	1	1.68	0.96	0.33
	Error	411.1	236	1.74		
Phoneme Identification	Grade	107.83	2	53.92	3.45	0.033
	Language	645.74	1	645.74	41.27	0.000
	Error	3132.07	236	27.65		
Word Identification	Grade	16622.59	2	8311.29	15.85	0.000
	Language			8693.99	16.58	0.000
	Error	12207.40	236	521.69		
Total score	Grade	20507.74	2	10253.87	16.59	0.000
	Language	4427.15	1	4427.15	7.161	0.008
	Error	145897.9	236	618.211		

Analysis of variance showed that there was no significant difference due to language on the letter identification task. However, there was a significant difference in the performance of the two groups on phoneme identification word identification and the total score ($F(1,236) = 41.27$; $p = 0.000$, $F(1,236) = 16.58$; $p = 0.000$ and $F(1,236) = 7.161$; $p = 0.01$ respectively).

There was a significant difference due to grade level for all the measures used in the test. Analysis of variance within the Chinyanja group showed that there was a significant difference between grades on all tasks: $F(2, 116) = 3.94$; $p = 0.02$ for letter identification;

$F(2,116) = 3.96; p = 0.02$ for phoneme identification; and $F(2,116) = 3.67; p = 0.03$; and $F(2,116) = 4.37; p = 0.02$ for the total score.

In the English group there was no significant difference in grade performance on letter identification and phoneme identification tasks. There was a significant difference in the performance of the children in this group between the grades on the total score and the word identification ($F(2,118) = 18.87; p = 0.000$ and $F(2,118) = 19.39; p = 0.000$ respectively).

CHAPTER FIVE: DISCUSSION

This chapter focuses on the discussion of the results. The item difficulty and discrimination indexes, the reliability, the distracter and the means and standard deviations of the two versions of the test are discussed in detail. The significance of the differences in the performance of the two groups of pupils is also discussed. The purposes of this study were to analyse the psychometric properties of the two versions of the ZAT – Reading Recognition and to find out if the two versions were comparable. The results showed that the two versions of the test had a wide variability of item difficulty and item discrimination. The two tests had too many easy items making the mean difficulty high.

5.1. Item Difficulty and Item Discrimination Indexes

The item difficulty and item discrimination of the two tests were almost the same for the letter identification items but there was a difference in the difficulty and discrimination indexes of the two versions of the tests for the phoneme identification and word identification items. The pupils in the Chinyanja condition performed poorly on the phoneme identification tasks. For example 91% of the children in the English group got item 83 right and only 72% of the Chinyanja group got a correct response for the same item. The poor performance might have been due to the fact that not all the children were native speakers of Chinyanja and the word *mvekeru* (which means sound) is not a word that children use in their everyday life. Although all the children who took

the test in Chinyanja said they spoke the language at home most of the time, not all of them were native Chinyanja speakers. A study done by Serpell (1970) on the Comprehension of Chi-Nyanja by Lusaka School children showed that native speakers of Chinyanja were superior to all other groups in the comprehension of Chinyanja on all the four measures of Chinyanja they were tested on. McAdams (1976) observed that where Chinyanja was the approved language only 49% of the pupils spoke Chinyanja at home. As phonological awareness was a fundamental aspect of reading children whose mother tongue is not Chinyanja may not do very well on tasks of this nature.

Item 95 was extremely difficult for the Nyanja group of children ($p = 0.03$; $D = 0.34$) compared to the English group ($p = 0.85$; $D = 0.18$). A critical look at the items shows that the picture for a rock (*thanthwe*) is not clear and it might have been quite difficult for the children to connect the picture to the word. The word is also rarely used.

The word recognition items were easier for the Chinyanja group of children. The pupils in the English group tended to find reading words more and more difficult as the test progressed. The last four items of the test were particularly difficult for this group of children but the discrimination index was good for this group. The Chinyanja items on word recognition (item 106 to 121) show very high discrimination values. This is because most of the children in the upper group could read all the word recognition items. These items do not provide any discrimination at all among the top group as

almost all the examinees in the high scoring group got all these items right. These items do not help us to differentiate between good readers and very good readers.

5.2. Reliability

The reliability of both versions of the tests was very high (Alpha Cronbach = 0.98 and 0.97 for the Chinyanja and English versions respectively). The high alpha was an indication that the test had high internal consistency. According to Coolican (2004) a test with internal consistency of 0.75 and above is acceptable and can be said to be internally stable. The two versions of the test were, therefore, both internally stable. The difference between the alpha values is also minimal (0.01) showing that the tests were very similar in this respect.

Item-total correlations for the letter identification tasks were generally low for the letter identification and phoneme identification tasks. For the Chinyanja versions of the test these correlations were in the range of 0.085 to 0.462. In the English version item-total correlations for these items were between 0.030 and 0.353. However, most of the items had significant correlations with the total score with the correlations being above 0.15 which is acceptable (Anastasi 1976, and Murphy and Davidshofer 1991).

The few items which did not meet the acceptable standard of correlation with the total score all correlated negatively with the total score. Item 69, 71, 72, 76, 75, 81 and 86 had negative item-total correlations in the Chinyanja test. Apart from item 81 all these

items had insignificant correlations with the total score. Item 81 was poorly correlated with the total score ($\alpha = 0.18$). This shows that this item does not measure what the rest of the test is measuring. The rest of items are so poorly correlated with the total score that they also do not measure the skill is supposed to be measured by the test. The poor correlations for these items could have been due to nervousness of some examinees as most of these items came quite early in the test.

Item 64, 69, 70 and 75 had negative correlations with the total score in the English version of the test. The correlations for these items were insignificant and the strange behaviour of these items can be attributed to anxiety and other confounding variables, such as experimenter effect, and anxiety of the pupils at the beginning of the test.

Although item 81 is exactly the same in the two versions of the test it behaves differently in the Chinyanja group. One obvious reason for this difference that it requires the Chinyanja group to identify a combination of letters that does not exist in the Chinyanja language. The letter X does not exist in Chinyanja and this is not used in writing. This means that children who are more familiar with Chinyanja than English may have difficulties recognizing a combination of X with another letter embedded in a word or string of letters. This appears to be clear indication of how writing system affects the way a child recognizes letters.

Items 102 to 121 were highly correlated with the total score in both versions of the test. In the Chinyanja version of the test the correlations were between 0.28 and 0.81 and in

the English test the correlations were between 0.45 and 0.80. These items had quite similar characteristics in terms of item-total correlations in the two versions of the test.

Pearson's correlations of the different tasks with each other and the total score in the Chinyanja test revealed that there was no significant correlation between the letter identification task and the phoneme identification task. This shows that the two tasks are not related to each other. Ideally tasks in a test should be interrelated as they are supposed to tap the same basic skills. There was a significant correlation between letter identification and word identification and the total score ($r = 0.199$; $P = 0.000$). These results appear to concur with Ryan, McNamara and Kenny (1977) that reading ability is significantly correlated to the ability to classify auditory sounds.

The fact that letter knowledge did not significantly correlate with phoneme identification can be accounted for by the fact that this group of children learnt the literacy skills in English. All the letters of the alphabet and their sounds were first learnt in English. This could have had an influence on the ability of the children in the Chinyanja group to segment the words in Chinyanja. The Chinyanja lessons the children received were also significantly fewer than English lessons. In some cases teachers do not teach the Chinyanja lessons because they do not speak the language themselves (Ohannessian 1976). This could have led to the poor ability of the children to transfer their letter knowledge to the Phoneme identification task.

The English version of the test showed significant correlations among all the tasks. This was an indication that all the tasks were measuring the same basic skill or the same construct. The correlations between phoneme identification and word recognition was higher than the correlation between letter identification and phoneme identification ($r = 0.41$; $p = 0.000$ and $r = 0.189$; $p = 0.019$ $N = 121$). This test conforms with the test theory rules that suggest that any tasks that are supposed to measure the same basic skills should be related to one another (Pedhazar and Schmelkin 1991).

Correlations of tasks were higher in the English version than in the Chinyanja version of the test. This shows that although the two tests are similar they are not exactly the same. These differences are due to the differences in the way the items in each of the test functions. The syntax in the instruction of the two versions is different as the two languages are totally different in the way the structure of the language is organized. The other reason that could have contributed to the discrepancy between the two tests is the fact that the vocabulary used in the phoneme identification task in the instructions for the Chinyanja test could have been too difficult for the children. The fact that instructions for literacy were initially given in English and the medium of instruction through which children learn all the other subjects, such as science and mathematics, was English meant that children were more exposed to how to recognize letters and words in English than Chinyanja. This could have resulted in the inability for the Chinyanja group to pay less attention to such things as word segmentation of Chinyanja words. The letter identification tasks were the same in both tests despite the facts that

some letters are not used in the Chinyanja writing system. As a result the children taking the test in Chinyanja had problems with letters such as X and Q.

Quality of Distracters

Distracter analysis results revealed that a few items had distracters that were too attractive, or confusing or both. According to classical test theory will distracters should be equally attractive to those examinees that do not know the correct answer (Murphy and Davidshofer 1991). Any distracter that is not chosen by any examinee is a poor distracter. A distracter that is too attractive and distracts even those examinees who know the answer is also a bad distracter.

In the Chinyanja version items 88, 91, 95, 98 each had a particularly bad distracter. In item 88 the distracter D was a very bad distracter as it distracted those who knew the answer and those who did not know the answer alike. 37.1 percent of the examinees chose this distracter while only 30.3 percent chose the correct answer. The rest of the examinees either chose B or C as the right answer. For items 91 and 98 the option A was a bad distracter as it was more attractive than the rest of the distracters put together.

In the English version of the test items 76, 78, 85, 89 and 98 each had a bad distracter. For items 78 and 85, option A was more attractive than the other two distracters put together. For item 76, C was a very attractive distracter and 12.4% of the examinees chose it as the correct answer where as none of the examinees chose B as the correct

answer. B is therefore a very bad distracter. For items 89 and 98 option D was a poor distracter as even some members from the high scoring group chose the distracter. For item 89, 18.2 percent of the examinees selected D as the correct response but only 9.1% selected the other two distracters. For item 98, 26.4% selected option D as the correct response and only 6.7% selected the other distracters .

For item 62 to 75 some of the distracters were not chosen at all in both versions of the test. This does not necessarily mean that these distracters were bad. These distracters were not selected because the items were very easy and straightforward and 92 to 100% of the participants got these items right.

Construct Validity

Factor analysis of the two versions of the test showed that both versions of the test were basically measuring one component. Although the loadings were slightly different on each task it was clear that the versions of the test were measuring the one construct. According to Pedhazzer and Schmekin, the absence of negative values in the loadings shows that the construct is a strong one.

Means, Standard Deviations and t-values.

The means for each grade in each group were computed for each task and for the total score. Scores were computed at grade level as achievement tests are often interpreted in terms of grade equivalent because these tests are usually used in a school setting. As a reading achievement test should provide information in the adequacy with which

essential content in reading skills are being taught it is important to look at the mean scores for each grade level for each group. A comparison of means for each grade showed that the only significant difference in the means of the two groups was at grade 4 level..

Grade 4 Means

The mean score for grade 4 (Chinyanja) was $N = 40$; $M = 70.2$; $SD = 30.33$. The mean for English group was $N = 41$; $M = 53.5$ $SD = 19.60$. The distribution of the scores in the two groups is very different. Ideally the distribution of the scores should produce a normal curve if the sample is taken from a cross-section of the same population (Anastasi, 1976). For these two groups the scores did not produce a normal curve. This shows that the test was not functioning as expected in the grade 4 pupils. In the English group the distribution of the scores were skewed to the left with most of the scores falling between 25 and 65. Only about one fourth of the pupils scored above 65 and the highest scored 105. It also shows that at Grade 4 level most of the children in the English group could not recognize even very simple words in English. In spite of having been given instructions in reading and writing from the first grade children at grade 4 level were unable to recognize words in English.

The distribution of scores in the Chinyanja group was bimodal showing that there was a group of children who could not recognize words at all while the other group was far ahead and could recognize words quite well. The mean score was 70.2 and the standard deviation was 30.33 showing that there was a lot of variance between individuals. This

bimodal distribution can be a result of some pupils picking up the rules of reading very quickly while others lag behind. McAdams (1976) observed that in the classroom there is usually a group of children who are far ahead of their counterpart in reading English. This can be attributed to the large numbers in the classroom from the first grade, which leaves the teacher with very little time to attend to all the children in the class. It also shows that a considerable number of children will have learnt how to read in Cinyanja by the time they are in grade 4.

The Means for the two groups were significantly different in favour of the Cinyanja group ($t(79) = 2.87; p = 0.005$). This shows that the Chinyanja group was far ahead of the English group in reading. This difference can be attributed to the fact that the Chinyanja orthography is regular and is easy to learn to read and write whereas the English writing system is more complex. Selecting pupils who said that they spoke English at home most of the time and were more comfortable taking the test in English for the English condition controlled the effects of language background. However, a few pupils spoke a language other than English and Nyanja but said they were more fluent in reading English than Cinyanja. As the Language background was quite controlled we can attribute this difference to differences in the orthography. These results concur with other studies that have been done in this area that suggest that in a regular orthography children learn to read faster than in a complex irregular orthography such as English (Wimmer and Goswami 1994; Landerl 2000 and Aro et al 2003).

Comparison of the means of different tasks of the test showed that there was no significant difference in the mean score for the letter- matching test at grade 4 level.

Letter knowledge is one aspect of reading that children learn quite early in school, and by the time they are in grade 4 almost all the children know the letters of the alphabet regardless of what language the alphabet is being learnt in. A study done by Serpell, Baker and Sonnenschein (2005) showed that by the time children were finishing Kindergarten, they identified 85% of the letter in their names. Since the two groups were taught initial literacy skills in English and the alphabet was initially introduced to them in English no difference in the performance of this task was expected.

On the phoneme identification task the children who took the test in English performed significantly better than the Chinyanja group ($t(79) = 2.97; p = 0.02$). On the other hand the Chinyanja group scored significantly higher ($t(79) = 3.57; p = 0.001$). The Chinyanja group was expected to score significantly higher on phoneme identification than the English group but the results were just the opposite of what was expected. Studies have shown that the ability to segment words is closely linked to the ability to recognize words (Vernon 1971). Goswami and Bryant (1990) studied the relationship between phonological awareness and word recognition among children and they found out that the progress children make in learning to read is critically dependent on the phonological skills of the children. Liberman et al, (1974) carried out a study on the relationship between phonemic segmentation and reading in children and their results showed that there was a relationship between phonemic segmentation and reading. Children who performed well on the phonemic segmentation task were good readers whereas those who did not do well on this task were poor readers. The results for the Chinyanja group were, therefore, surprising, as they showed no relationship at all between phoneme identification and word recognition.

One reason for the poor performance on this task was that the instructions in the English version were very simple and easy to follow. In the Chinyanja language the instructions seemed to have been more difficult for the children to understand because the word ‘sound’ (*mvekeru*) is not a word that is commonly used in Chinyanja. This clearly illustrates how translation problems at the vocabulary level can affect the psychometric properties of a translated test. Some of the pictures used for this task had lexis that the children at grade 4 level probably have never heard of. A good example of this is the picture of a tent (*hema*) in item 84. Most children in Lusaka know a tent by its loaned word from English *tenti* and even if the examiner points to the picture and says ‘*hema*’ the child may not be able to connect the word *hema* to the picture of the tent as it is not in his or her vocabulary.

Williams (1998) also observed that the standard Chinyanja that is used in the school books is often not the Chinyanja that children in the area speak. For example in the standard Chinyanja a crocodile is *ng’ona* but many Lusaka children know it as *ngwen*, (the Chinsenga word for crocodile). Some of the children would actually try to correct the examiner if she said that the stimulus picture in item 96 was a picture of *ng’ona*, some children would promptly said that it was not a *ng’ona* but a *ng’wena* (personal experience). This clearly shows the differences between the standard Chinyanja and the Chinyanja that is spoken by the children.

Grade 5 Means

The mean scores for the letter matching task were not significantly different (N =39; M= 19.26 and N = 40; M = 18.9, for English and Nyanja respectively. The same reasons as

for the grade 4 level can explain the fact that the two means are very similar. There was a marked difference in the performance of the two groups on the phoneme identification task. The English group performed far much better than the Chinyanja group ($t(77) = -5.92$; $p = 0.000$). The poor performance by the Chinyanja group is more apparent than at grade 4 level. This disparity may be accounted for by the fact that the children in the English group who had a slow start are now catching up with their friends. As the English group becomes more fluent in the language their ability to detect the presence sound segments in the spoken words improves. At grade 5 level the language that is used most of the time in the classroom is English as pupils are expected to have attained a sufficient amount of English to help them get through the school day. On the other hand as a child moves up into grade 5 and 6 the Chinyanja lessons are reduced and teachers concentrate on other subjects such as mathematics and social studies. (McAdams 1976 and Ohannessian (1976). Less attention is paid to the teaching of local languages as, although they are examinable at grade 7, they are not counted in the qualifications for Grade 8. The Chinyanja group therefore, has less time practicing how to read the language. In addition to the problems of the vocabulary in both the instructions and some of the stimulus pictures the grade 5 pupils also have to contend with lack of practice.

The means for the total score of the two groups were not significantly different. However, the distribution of scores in the two groups was quite different. In the Chinyanja group, the distribution of the scores shows that at grade 5 there were still a few pupils who could not recognize words and also performed quite poorly on the

phoneme identification task. This shows that although Chinyanja is easy to learn to read and write some children still cannot read at grade 5. These are the children who could have reading problems. This finding concurs with studies done by other researchers that suggest that children who are unable to read in the first years of school often fall far behind their contemporaries (Vernon: 1971, Nunes 2000).

The distribution of scores in the English group was slightly skewed to the left with most of the scores falling below the mean and very few scores at the tail on the right. This shows that when children learn to read in English only a few children will have learnt how to recognize words quite well whereas, many of them will still be struggling to recognize words.

Grade 6 Means

At grade 6 level the results for letter matching were similar to grades 4 and 5. There was almost no difference in the mean scores, $N = 41$; $M = 19.26$; $SD = 0.98$ for the English group and $N = 39$; $M = 19.25$; $SD = 1.98$ for the Chinyanja group. These results were expected, as there was no difference in the two subtests. Children at Grade 6 level are also expected to have mastered their alphabet and only those with serious learning difficulties fail to recognize letters at this level.

The results for phoneme identification showed that the English group of children performed better on the task but the difference was not significant. The Chinyanja group scored significantly higher on the word recognition task but the difference was not significantly higher on the word recognition task. The Chinyanja group of children still

had problems in detecting the segments in a word at grade 6. However, the Chinyanja group was still ahead of the English group in terms of the number of words that they could recognize. Studies have shown that children learning to read in a regular orthography can read virtually any word once they have mastered the rules of reading (Wimmer and Goswami, 1994, and Aro et al, 2003). As the Chinyanja writing system is very regular children who learnt how to read in Grade 4 can read all the words in Grade 6.

Overall Performance

A look at the overall performance of the two groups showed that in the Chinyanja group there was a significant amount of variance due to the effect of grade ($F(12,116) = 3.94$; $p = 0.021$), that is, the higher the grade the better the performance was. Although children are expected to recognize all the letters of the alphabet by the time they are in grade 4 children who say they can recognize letters in Chinyanja may still find it difficult to recognize letters of the alphabet that are absent in the Chinyanja language such as Q and X. By the time they are in grade 5 this problem disappears.

In the group of children who took the test in English the variance in the letter and phoneme-matching tasks due to grade was insignificant. By grade 4 children have learnt to recognize letters of the alphabet and are already paying attention to the segments in words. Bruce (1964) found out that children below the mental age of 7 were unable to analyse words correctly and children above the age of 7 had greater success in analysing or segmenting words. As the children in our sample were all above 7 years it was expected that they would find the two tasks easy.

In the Chinyanja group grade was still a factor that accounted for differences in performance on the phoneme identification task ($F(2, 116) = 3.91; p = 0.22$). Most of this variance was due to the differences between grades 5 and 6. In fact the grade 5 pupils mean score was less than that for grade 4. These results could indicate that by grade 6 many of the pupils have grasped enough of the standard Chinyanja that they are able to understand the instructions quite well. Some words that were not in their vocabulary in grade 4 are now familiar to them and they can easily make connections between the words and the stimulus pictures.

Variance due to grade in both languages was significant; $F(2, 116) = 3.67; p = 0.03$ for Chinyanja and $F(2, 118) = 19.39; p = 0.000$. There was more variance between grades in English group than in the Chinyanja group. Within the English group there was no significant difference in the performance of grade 4 and grade 5. Most of the variance due to grade was accounted for by the variance in mean scores between grades 4 and 6 and between grades 6 and 5). This shows that by grade 5 children in the English group still had problems in recognizing words. These differences could also be explained by the complexity and irregularity of the English writing system. The irregularity in the writing system makes it difficult for children to recognize words that they have not come across before. This is because the English writing system has so many rules and it has more than one way of reading the same combination of vowels. For example, the words *bead* and *dead* the *ea* is pronounced differently. In such instances a child has to know the word in order to read it. (Aro, unpublished). Words like *eat*, *good* and *houses* are easy to read for the grade 4s because they come across them in their readers all the time but words like *reign*, *heroine* and *limousine* are not words that children see in print or

hear regularly. As the rules that apply to reading English are so many it takes children number of years to be able to read at an acceptable level. Words that are not in the children's vocabulary are most difficult to read (Vernon 1971). The large variance due to grade shows that by the time children are in grade 6 they have learnt many of the rules and exceptional in reading English.

There were significant differences in variance of the total score between grades in both groups ($F(2,116) = 4.37; p = 0.02$ for the Chinyanja group and $F(2,118) = 18.87; p = 0.000$). This was expected because as children graduate from one grade to another their ability to grasp things becomes more. Instruction also had an effect. Williams (1998) in his study on literacy in Malawi and Zambia also found similar results. He found out that reading performance improved as the children moved up the grades in both languages (Chinyanja and English).

Analysis of variance revealed that the between group variance on the letter matching task was insignificant but there were significant differences due to language on the phoneme identification task and word identification tasks ($F(1,236) = 41.27; p = 0.000$ and $F(1,236) = 16.58; p = 0.000$). The between language groups variance in the mean total score was significant at $F(1,236) = 7.16; p = 0.008$.

The difference in the overall variance on phoneme identification can be attributed to the reasons that have already been mentioned in this discussion. The main reason is that children who took the test in Chinyanja may have misunderstood the instructions as the word used in Chinyanja as the equivalent word for the English word sound (*mvekeru*)

may have been too technical for the children to understand. The children in the sample all went through the English immersion program. These children had fewer lessons in Chinyanja than English. Studies have shown that the more a child is exposed to print the more fluent a reader that child will be. Ehri (1979) also observed that the amount of training or practice provided for children has an effect on their ability to analyse sounds. The effect of practice was particularly seen in the phoneme identification task where the children in the English group performed better than the Chinyanja group. However, when we look at the mean total score the children in the Chinyanja condition were superior to those in the English condition in spite of the fact that the both groups had more training and practice in English.

Williams (1998) also observed that in Zambia the children who said they were comfortable reading in English often could not read at all. In his study children who preferred to read in Chicewa in Malawi said they preferred to read in that language because it was easier to read. The high overall mean score by the Chinyanja group can, therefore, be explained by the transparency in the writing system of the Chinyanja orthography.

The overall distribution of scores for Grades 4, 5 and 6 was relatively normal for the English group although there were fewer scores on the right. The Chinyanja distribution was bimodal with a considerable number of children falling two standard deviations below the mean and a large group of children falling above the 50th percentile. There were very few scores around the mean. The differences in the distributions show how

the same test given in two languages to the same populations produced very different kinds of distribution of scores. The large differences in the way the scores are distributed shows that the two versions of the test did not function in exactly the same way in the two samples.

CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS

Summary of Results

The purpose of this study was to analyse the psychometric properties of the two versions of the ZAT-RR. Results revealed that both versions of the test had items with a wide variety of difficulty although the Chinyanja version was easier than English version. A few items had negative item discrimination indexes showing that they might have been poorly constructed. However, all the items with negative discriminations in both versions were very easy items that came quite early in the test. These negative discriminations were also of very low value ($D = 0.03 - 0.06$) these negative discrimination were due to confounding variables such as examined anxiety, and experimenter situational effect. Some item were easier in one language but more difficulty for the other group. Items 113 to 121 were more difficult for the English group than the Chinyanja group of children.

The two versions of ZAT-RR showed very high internal consistency with alpha values of (0.98 and 0.97). This showed that the test was reliable and internally stable. The mean total scores at grade 4 level were significantly different indicating that children found Chinyanja easier to read and learn it quite early. There was no significant difference between the mean total scores for grade 5 and 6.

Factor analysis revealed that both tests were testing one basic skill. The loadings on this factor for all the tasks in the subtest were quite high. Generally there were high loadings

on the word recognition task 0.798 and 0.812 for Chinyanja and English respectively. Letter recognition had the lowest loading on the factor being measured in both tests.

Conclusion

It can be concluded that the two versions of the test, although not exactly equivalent are quite similar in their psychometric properties. The major difference was in the phoneme identification task. With some improvements to the present Chinyanja test, the tests can be comparable. This study has also clearly shown how difficult it is to translate a reading test from one language into another. The study has also illustrated how writing systems affects the psychometric properties of a reading test. However the test is highly reliable and it is valid. It can therefore, be used once the items that have very poor psychometric properties have been replaced.

6.1 Recommendations

1. The instructions for items 82 to 101 in the Chinyanja test of the ZAT-RR should be revised so that they are clear and easily understood by primary school children.
2. Item 95 in the Chinyanja version has to be revised. The picture of a rock should be replaced with something that looks more like a rock.
3. Item 81 in the Chinyanja version should be replaced, as it does not contribute to the effectiveness of the test.

on the word recognition task 0.798 and 0.812 for Chinyanja and English respectively.

Letter recognition had the lowest loading on the factor being measured in both tests.

6.2 Conclusion

It can be concluded that the two versions of the test, although not exactly equivalent are quite similar in their psychometric properties. The major difference was in the phoneme identification task. With some improvements to the present Chinyanja test, the tests can be comparable. This study has also clearly shown how difficult it is to translate a reading test from one language into another. The study has also illustrated how writing systems affects the psychometric properties of a reading test. However the test is highly reliable and it is valid. It can therefore, be used once the items that have very poor psychometric properties have been replaced.

6.3 Recommendations

1. The instructions for items 82 to 101 in the Chinyanja test of the ZAT-RR should be revised so that they are clear and easily understood by primary school children.
2. Item 95 in the Chinyanja version has to be revised. The picture of a rock should be replaced with something that looks more like a rock.
3. Item 81 in the Chinyanja version should be replaced, as it does not contribute to the effectiveness of the test.

4. Some items in both versions of the test need to be repositioned. Items that are more difficulty but come earlier in the test should be moved. For example item 91 is easier in both versions than item 90 and item 92 is more difficult than items 92, 93, 94 and 95. For the Chinyanja test all the items from item 82 to 101 need to be reordered as easier items come later than some difficulty items. Timing of the word recognition test for the Chinyanja test would help this part for the test to be as difficulty as the English part of the test.
5. English and Chinyanja are totally different languages. Some letters are hardly used at the beginning of words in Chinyanja. For example in Chinyanja very few words begin with the letter 'H' and very few names of objects begin with 'L'. It would, therefore, be advisable to replace pictures that start with these letters.
6. Further research is required to find out if translated tests are comparable.

REFERENCES

- dam, M.J. (1990) *Beginning to Read: Thinking and Learning about Print*. Cambridge: MIT Press.
- iken, R. (1994) *Psychological Testing and Assessment*. Boston: Allyn & Bacon
- nsastasi, A.(1988) (6th ed) *Psychological Testing*. New York: Macmillan Publishing Company.
- nsastasi, A. (1976)(4th ed.) *Psychological Testing*. New York: MacMillian Publishing Company.
- ro, M. (2003) "Learning to Read: English in Comparison to Six More Regular Orthographies". *Applied Psycholinguistics*, 24, 621 – 635.
- ro, M. (2004) "Learning to Read: The Effect of Orthography." Unpublished PhD Thesis Jyvaskyla.
- ro, M., Tolvanen, A., Poikkeus, A., and Lyytinen H. (2004) "The Development of Reading and Spelling Skills, and the Predictors of Accuracy and Fluency: An intensive follow-up in a Transparent Orthography (Finnish)". Jyvaskyla: Nilo Maki Institute
- uce, D. (1964) "Analysis of Word Sounds by Young Children." *British Journal of Educational Psychology*, 34, 158- 170.
- olican, H. (2004) *Research Methods and Statistics in Psychology*. Bistol: Hodder and Stoughton Ltd.
- ockett, D. (2004) "Assessment Practices Used in Bilingual and LEP Students." www.GASPnet.org.
- onbach, L.J (1949) *Essentials of Psychological Testing*. New York Harper and Brothers Publishers.
- fior, S., Martos, F. and Cary, L. (2002) Differences in Reading Acquisition Development in Two Shallow Orthographies: Portuguese and Spanish. *Applied Linguistics*, 23, 135-148.
- i, L (1979) "Linguistic Insight: Threshold of Reading Acquisition". In R.G. Waller and G.E Mackinnon (Eds). *Reading Research: Advances in Theory and Practice*, New York: Academic Press.
- i, L.C., Nunes, S. R., Willows, D.M. Schuster, B.V. Yaghoub- Zadeh, Z., and Shanahan, T. (2001) "Phonemic Awareness Instruction helps Children Learn to Read: Evidence from the National Reading Panel's Meta-analysis". *Reading Research Quarterly*. 36, 250-287.

- Frith, U. (1980) *Cognitive Processes in Reading*. London: Academic Press.
- Goswami, U. and Bryant, P (1990) *Phonological Skills and Learning to Read*. Exeter: Erlbaum.
- Grigorenko, E. L., Jarvin, L., Kaani, B., Kapungulya, P. P., Kwiatskoski, J. and Sternberg, R. J. (2003). Effects of Interventions against Parasitic Illnesses on a Test of Cognitive Skills Administered in Zambia. Unpublished manuscript.
- Kashoki, M. E. (1990) *The Language Factor in Zambia*. Lusaka: Kenneth Kaunda Foundation.
- Kehoe, J. (1995) "Basic Item Analysis for Multiple-Choice Tests. *Practical Assessment, Research and Evaluation*, 4, 10, 130-145.
- Kelly, M. J. and Kanyika, J. (2000) *Learning Achievement at the Middle Basic Level: Summary Report on Zambia's Assessment Project 1999*. Lusaka: Zambia Ministry of Education.
- Landerl, K. (2000). "Influences of Orthographic Consistency and reading instruction on the Development of Nonword Reading Skills." *European Educational Psychology*, 15, 239-257.
- Lyytinen, H., Aro, M., Halopainen, L., Leiwo, M., Lyytinen, P. and Tolvnen, A. "Children's Language Development and Reading Acquisition in a Highly Transparent Language." In R. M. Joshi and P. G. Aaron(Eds) *Handbook of Orthography and Literacy*. (in press).
- Liberman, I., Y., and Shankweiler D. (1977) "Speech, the Alphabet and Teaching to Read." In L.B. Resnick and P.A Weaver (Eds) *Theory and Practice of Early Reading*. Hillsdale, N.J: Lawrence Erlbaum Association.
- Liberman, I.Y., Shankweiler, D., Fischer, F.W., and Carter, B. (1974) "Explicit Phoneme and Syllable segmentation in Young Children." *Journal of Experimental Child Psychology*, 18, 201 - 212.
- Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Kulju, P., Laakso, M. L., Leiwo, M., Leppanen, P., Lyytinen, P., Poikkeus, P., Richardson, U., Torppa, M. and Viholainen, H. (2004) "Early Development of Children at Familiar Risk for Dyslexia Follow-up from Birth to School Age." *Dyslexia* 10, 146 - 178.
- Machungwa, P. and Ali, M.R. "Psychological Testing in Zambia." Unpublished.
- Mann, V. and Wimmer H. (2002) "Phoneme Awareness and Pathways into Literacy: A comparison of German and American Children." *Reading and Writing: An interdisciplinary Journal*, 15, 653 - 682
- McAdam, B. (1976) "The New Zambia Primary Course." In S. Ohannessian and M.E Kashoki (Eds) *Language in Zambia*, London: International Institute of African Studies.
- Ministry of Education (1975) *Zambian Languages Orthography*. Lusaka: Ministry of Education.

- Murphy, R. K. and Davidshofer, O. C. (1991) *Psychological Testing: Principles and Applications*. New Jersey: Apprentice- Hall Inc.
- Nunes, T. (2000) Learning to Read: An Integrated View from Research and Practice. www.literacytrust.org.
- Ohannessian, S. (1976) "The Teaching of Zambian Languages and the Preparation of Teachers for Language Teaching in Primary Schools." In S. Ohannessian and M. E Kashoki (Eds). *Language in Zambia* London: International Institute of African Studies.
- Pedhazur, E. J. and Schmelkin, P.L. (1991) *Measurement, Design and Analysis :An Integrated Approach*. New Jersey:Lawrence Erlbaum Association Publishers.
- Rayner K. and Pollatsek, A. (1989) *The Psychology of Reading*. New Jersey: Price-Hall Engelway.
- Ryan, E.B., McNamara, S.R., and Kenny, M. (1977) "Lexical Awareness and Reading Performance Among Beginning Readers." *Journal of Reading Behaviour*, 9, 399-400.
- Sechrest, L., Fay, T.L., and Zaidi S. M H. (1972) "Problems of Translation in Cross-Cultural Research". *Journal of Cross-Cultural Psychology*, 3,1, 41-56.
- Sekeleti, C. (1988) "The Medium of Instruction in Zambian Schools". *Psychological Studies*, 8, 38.
- Serpell, R. (1970) "Chi-Nyanja Comprehension in Lusaka School Children: A field Experiment in Second Language Learning." *H.D.R.U. Reports*, 16, 1970.
- Serpell, R. Baker, L. and Sonnenschein, S. (2005) *Becoming Literate in the City: The Baltimore Early Childhood Project*. New York: Cambridge University Press.
- Seymour, H.P., Aro, M. and Erskine, J. (2003) " Foundation Literacy Acquisition in European Orthographies." *British Journal of Psychology* 94, 143-174.
- Seymour, H. P. (1986) *A Cognitive Analysis of Dyslexia*. London: Routeledge & Kegan Paul.
- Sharma, R. (1973) "The Reading Skills of Grade III Children." Lusaka:Ministry of General Education.
- Singleton, C. and Vincent, D. (2004) "Assessing Literacy: Current Challenges and Issues." *Journal of Research in Reading*, 27, 2, 113-117.
- Trieman, R.(1993) *Beginning to spell*. New York: Oxford University Press.
- Wernon, M.D (1971). *Reading and its Difficulties: A Phonological Study*. London: Cambridge University Press.

Williams E. (1998) "Investigating Bilingual Literacy: Evidence from Malawi and Zambia." *Education Research*, 24, 99-130.

Wimmer, H. and Goswami, U., (1994) "The Influence of Orthographic Consistency on Reading Development: Word recognition in English and German Children." *Cognition*, 51, 19-103.

Appendix A: Consent Form used to seek permission from participants' parents

Parental Consent Form

Name of Child:
Age: Sex: Admission number:
School:
Zone / District:
School code / EMIS number:

It has been decided to establish a program to understand how children in Zambia learn. Our goal is to improve the services provided to school children in school. The University of Zambia, in collaboration with colleagues from Yale University in the United States, is conducting a survey of learning abilities and difficulties. We are talking both to children who currently attend school and to those who don't. By participating in this study, you are not committing yourself or your child to go to school.

Your village has been selected to take part in this program and we request your permission to allow your child to take part. Should you agree, your child will be asked to take tests that will help us understand how s/he learns best. The tasks are similar to things that children do in school and to tasks that they need to accomplish in their daily lives in your village.

We will also ask your child to let us take four photographs of his or her head and hands. We will take a fifth portrait photograph that we will give to your child to keep.

In addition, we will measure and weigh your child, take a picture of your child, and collect a sample of saliva. The saliva collection is done by using a brush [show brush] that we sweep on the inside of the cheek. The purpose of collecting the saliva and photographs is to enable us conduct a thorough analysis for learning disabilities. There are no expected risks to your child's participation in this program.

All of your child's answers will be confidential and will not be shared with his / her teachers or the village chiefs. We will use the information your child provides to improve the way students are instructed in schools in Zambia.

If your child is identified as having difficulties with learning, we will return to you and ask for your permission to collect additional information from you and your child. By signing this document, you are also agreeing to allow us to contact you again about additional participation.

Declaration by parent / guardian:

I give my consent for my child to take part in the student learning program. I understand that I am free to withdraw my child from the study at any time.

Name of parent:.....Signature / Mark:

For any queries, please contact the University of Zambia, Research Ethics Committee P. O. Box 50110 Ridgeway Campus, Lusaka. Tel: 260-1-256067 Telefax: 260-1-250753 E-mail: unzarec@zamtel.zm

Appendix B: Sample of items in the ZAT RR

5. Point to the "C" at the top of the page.

Say: **Find one like this—down here.** Point to the area with the other letters.

Say: **Point to it.**

C

Q	U
D	C

#23

23. Point to the picture of the horse at the top of the page.

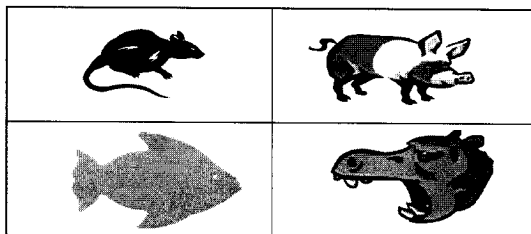
Say: **This is a picture of a horse.** Point to the area with the other pictures.

Say: **Which of these pictures begins with the same sound as this picture? Point to the word that begins with the same sound.**

Answer: Bottom Right

ZAT-RR-English

46



ZAT-RR-English

47

45. Point to each row of words.

Say: **This page has four words on it. Read each word aloud. Go slowly from one word to the next.**

outside	fishing
town	smile

Appendix C: Item analysis results

ITEM	it62	it63	it64	it65	it66	it67	it68	it69	it70	it71	it72		
L2 P	0.95	1	1	0.98	0.97	0.95	1	0.97	0.98	0.76	0.95		
L1 P	0.86	0.94	0.98	0.94	0.88	1	0.98	0.94	0.98	0.91	0.94		
L2 D	0.09	0	0	0.03	0.06	0.03	0	0	-0.03	0.18	0.03		
L1 D	0.16	-0.06	-0.03	0.06	0.19	0	0.03	0.06	0.03	0.06	0.06		
ITEM	it73	it74	it75	it76	it77	it78	it79	it80	it81	it82	it83		
L2 P	0.97	1	0.95	0.91	0.95	0.82	0.95	0.89	0.92	0.91	0.92		
L1 P	0.97	0.98	0.98	0.92	0.98	0.88	0.97	0.84	0.94	0.72	0.84		
L2 D	0	0	-0.03	0.18	0.03	0.18	0.03	0.09	0.09	0.06	0.09		
L1 D	0	-0.03	0.03	0.03	0.03	0.19	0	0.19	0.13	0.25	0.13		
ITEM	it84	it85	it86	it87	it88	it89	it90	it91	it92	it93	it94		
L2 P	0.7	0.73	0.89	0.89	0.64	0.71	0.85	0.45	0.76	0.73	0.85		
L1 P	0.56	0.83	0.77	0.75	0.53	0.63	0.8	0.36	0.59	0.48	0.58		
L2 D	0.12	0.24	0.21	0.15	0.3	0.33	0.24	0.36	0	0.42	0.18		
L1 D	0.47	0.28	0.03	0.13	0.28	0.25	0.28	0.34	0.37	0.28	0.53		
ITEM	it95	it96	it97	it98	it99	it100	it101	it102a	it102b	it102c	it102d		
L2 P	0.85	0.91	0.91	0.67	0.88	0.92	0.92	0.77	0.74	0.91	0.76		
L1 P	0.03	0.72	0.88	0.69	0.84	0.75	0.92	0.83	0.66	0.69	0.67		
L2 D	0.18	0.18	0.3	0.3	0.18	0.1	0.16	0.18	0.46	0.18	0.3		
L1 D	0.34	0.18	0.13	0	0.29	0.44	0.16	0.25	0.63	0.5	0.59		
ITEM	it103a	it103b	it103c	it103d	it104a	it104b	it104c	it104d	it105a	it105b	it105c	it105d	it105d
L2 P	0.83	0.74	0.62	0.5	0.76	0.71	0.55	0.48	0.42	0.5	0.56	0.55	
L1 P	0.73	0.75	0.66	0.64	0.86	0.72	0.58	0.58	0.53	0.58	0.52	0.56	

L2D	0.34	0.52	0.7	0.94	0.48	0.58	0.91	0.97	0.79	0.88	0.82	0.84
L1D	0.53	0.5	0.69	0.72	0.28	0.56	0.78	0.81	0.94	0.94	0.97	0.88

ITEM	it106a	it106b	it106c	it106d	it107a	it107b	it107c	it107d	it108a	it108b	it108c	it108d
L2P	0.59	0.64	0.64	0.35	0.64	0.65	0.53	0.53	0.5	0.52	0.5	0.42
L1P	0.44	0.48	0.52	0.5	0.47	0.47	0.48	0.5	0.45	0.52	0.55	0.56
L2D	0.82	0.67	0.73	0.7	0.91	0.7	0.88	0.76	0.94	0.97	0.94	0.84
L1D	0.69	0.97	0.97	1	0.88	0.75	0.91	0.88	0.9	0.97	0.91	0.88

ITEM	it109a	it109b	it109c	it109d	it110a	it110b	it110c	it110d	it111a	it111b	it111c	it111d
L2P	0.47	0.5	0.41	0.73	0.38	0.48	0.41	0.44	0.58	0.39	0.27	0.38
L1P	0.52	0.48	0.5	0.48	0.53	0.52	0.45	0.45	0.52	0.5	0.48	0.5
L2D	0.76	0.94	0.76	0.79	0.76	0.97	0.76	0.88	0.84	0.67	0.55	0.76
L1D	0.97	0.97	0.94	0.84	0.94	0.97	0.91	0.84	0.97	1	0.97	0.88

ITEM	it112a	it112b	it112c	it112d	it113a	it113b	it113c	it113d	it114a	it114b	it114c	it114d
L2P	0.73	0.44	0.26	0.47	0.52	0.35	0.33	0.29	0.38	0.35	0.23	0.14
L1P	0.47	0.52	0.47	0.52	0.56	0.45	0.52	0.52	0.52	0.45	0.52	0.56
L2D	0.55	0.88	0.52	0.94	0.91	0.64	0.67	0.52	0.7	0.7	0.45	0.28
L1D	0.94	0.91	0.94	0.94	0.82	0.88	0.97	0.84	0.94	0.82	0.91	0.88

ITEM	it115a	it115b	it115c	it115d	it116	it116b	it116c	it116d	it117a	it117b	it117c	it117d
L2 P	0.39	0.44	0.35	0.23	0.44	0.38	0.35	0.3	0.18	0.32	0.36	0.09
L1 P	0.5	0.55	0.5	0.5	0.52	0.53	0.45	0.41	0.59	0.39	0.41	0.52
L2 D	0.79	0.76	0.7	0.46	0.88	0.76	0.7	0.6	0.36	0.64	0.73	0.18
L1 D	0.94	0.84	1	0.94	0.97	0.91	0.91	0.82	0.63	0.78	0.83	0.91

ITEM	it118a	it118b	it118c	it118d	it119a	it119b	it119c	it119d
L2 P	0.26	0.27	0.2	0.15	0.33	0.35	0.38	0.26
L1 P	0.53	0.53	0.52	0.53	0.48	0.45	0.45	0.48
L2 D	0.45	0.55	0.39	0.3	0.67	0.7	0.76	0.52
L1 D	0.94	0.94	0.91	0.88	0.97	0.91	0.84	0.81

ITEM	it120a	it120b	it120c	it120d	it121a	it121b	it121c	it121d
L2 P	0.32	0.32	0.15	0.14	0.35	0.26	0.38	0.17
L1 P	0.5	0.48	0.52	0.48	0.44	0.52	0.52	0.47
L2 D	0.64	0.55	0.3	0.21	0.7	0.55	0.76	0.33
L1 D	0.94	0.84	0.91	0.72	0.88	0.84	0.84	0.81

L1 = Chinyanja

L2 = English

P = item difficulty index

D = Item discrimination index